17641

NATIONAL LIBRARY
OTTAWA

CANADA

BIBLIOTHÈQUE NATIONALE
OTTAWA

NAME OF AUTHOR....R. W. Olafson.....

TITLE OF THESIS....Physico-Chemical and Sequence Studies on Streptomyces griseus trypsin

UNIVERSITY...of Alberta.....

DEGREE FOR WHICH THESIS WAS PRESENTED....Ph.D.....

YEAR THIS DEGREE GRANTED.....1973.....

Permission is hereby granted to THE NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(Signed).....

PERMANENT ADDRESS:

1926 Grawley St.
Vancouver 6, B.C.
.....

DATED....June 13.....1973

NL-91 (10-68)

THE UNIVERSITY OF ALBERTA

PHYSICO-CHEMICAL AND SEQUENCE STUDIES ON *STREPTOMYCES GRISEUS* TRYPSIN

by

© ROBERT WALTER OLAFSON

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOCHEMISTRY

EDMONTON, ALBERTA

FALL, 1973

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled "Physico-chemical and Sequence Studies on Streptomyces Griseus Trypsin" submitted by Robert Walter Olafson in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

L. B. Smillie
Supervisor

W. Ingram

Cyril M. Kay

J. Thompson

Hans Neurath
External Examiner

Date June 11, 1973.

ABSTRACT

Serine proteases are a group of proteolytic enzymes, widely distributed in nature and all inhibited by the organophosphorous compound, DFP. Historically, they were first recognized among the vertebrate digestive enzymes originating in the pancreas, but have more recently been found free in the blood, as products of the intestinal mucosa and throughout the invertebrate and microbial worlds.

The past decade has witnessed considerable research into the structure and mechanism of catalysis of both eucaryotic and procaryotic serine proteases. One such protease obtained from the latter group has been isolated in an active form from Pronase, an extracellular filtrate of the organism Streptomyces griseus. Pronase contains both exo- and endopeptidases along with material lacking proteolytic activity. Streptomyces griseus trypsin (SGT) is an endopeptidase which has been isolated from Pronase by fractionation on CM-Sephadex followed by SE-Sephadex, using a linear gradient of pyridine-acetic acid buffer, pH 5.0, with both ion exchange systems.

SGT, prepared in the above manner, has been shown to be remarkably similar to bovine trypsin by both physico-chemical and structural criteria. For example, stability studies indicated that SGT, like bovine trypsin,

was more stable in the presence of calcium ion than in its absence. Kinetic experiments, using BAEE as substrate, have similarly shown that SGT has a $K_{m(app)}$ (8×10^{-6}M) that is not unlike trypsin (10^{-5}M). SGT is also inhibited by TLCK and DFP—both known inhibitors of trypsin activity. TLCK was demonstrated to specifically alkylate the active site histidine-57 in SGT—again analogous with trypsin. Furthermore, the bacterial enzyme has been found to exhibit a marked kinetic isotope effect. Using the substrate BAEE dissolved in deuterium oxide, the reaction rate was observed to be 1/3 that observed in water. This finding has been interpreted by other workers to indicate the presence of a rate limiting proton transfer in the catalytic mechanism of serine proteases. SGT also exhibits a narrow 'trypsin-like' specificity for hydrolysis of the insulin B chain, cleaving the polypeptide at lysine and arginine residues exclusively.

Extensive investigations directed towards elucidation of the role of the NH_2-terminal α-amino group in SGT gave no indication that the enzyme required a charged α-amino group for activity. Furthermore, no conformational changes were exhibited on titration in the alkaline pH range, as estimated by ORD and CD studies. This is in contrast with other serine proteases which appear to require such a charged group, but in agreement with results obtained for α-lytic protease of Myxobacter

495, which does not.

The molecular weight of SGT, as estimated by ultracentrifugation studies and corroborated by sequence analysis, was found to be 22,900. The molecular weight of bovine trypsin is 24,000.

The complete primary sequence of SGT has been elucidated and a speculative ribbon diagram of the molecule constructed, using the chymotrypsin backbone as a reference structure. In no case did a hydrophilic group in the SGT sequence replace an internal hydrophobic residue of chymotrypsin. The degree of homology between the buried residues of chymotrypsin and corresponding residues in SGT was remarkably high. Using this structure and the known sequence of SGT, a comparison of the salient features of the reported structure of bovine trypsin was also made. The data compiled are virtually unanimous in their indication of the high degree of similarity between SGT and bovine trypsin, strongly suggesting that both enzymes evolved from a common ancestral protein.

ACKNOWLEDGEMENTS

The author wishes to express his appreciation to Professor L. B. Smillie in whose laboratory the investigations in this thesis were conducted and for his continued advice during the development of this work.

The author also expresses his sincere thanks to Dr. L. Jurasek whose constant interest and advise proved invaluable throughout the progress of the research.

The author is indebted to Mr. J. M. Nattriss for his skilled technical assistance during amino acid analyses.

The author further wishes to extend his appreciation to Mr. M. R. Carpenter for his skilled technical assistance in operation of the Beckman Sequencer.

The author expresses his gratitude to Professor C. M. Kay and members of his laboratory; Dr. W. D. McCubbin for valuable advice regarding sedimentation equilibrium molecular weight determinations and spectropolarimetric studies and Mr. K. Oikawa for his technical advice and assistance during ORD and CD studies.

The author is indebted to Mr. M. Aarbo for technical assistance rendered during sedimentation equilibrium studies.

Finally, the author wishes to express his indebtedness to the Medical Research Council of Canada for both financial support of the research described in this thesis and personal support in the form of a Medical

Research Council Studentship for the years 1969 to 1972. He also acknowledges financial support in the form of a University of Alberta graduate teaching assistantship for academic years 1972 to 1973.

TABLE OF CONTENTS

CHAPTER		PAGE
	Abstract	iv
	Acknowledgements	vii
	List of Tables	xv
	List of Figures	xviii
	List of Abbreviations	xxii
I.	INTRODUCTION	1
II.	GENERAL MATERIALS AND METHODS	46
	1. MATERIALS	46
	A. General Chemicals and Solvents	46
	B. Enzymes	47
	2. METHODS	48
	A. Column Chromatography Procedures for Preparation of SGT from Pronase	48
	(1) Ion Exchange Gel Chromatography	48
	(2) Bio-Rex-70 Chromatography	49
	B. Column Chromatography Procedures for Purification of Peptides	50
	(1) Cyanogen Bromide Peptides Separated on Sephadex G-75 Superfine.	50
	(2) Sephadex G-25 Fine and G-50 Fine Purifications of Tryptic Digests of Performic Acid Oxidized SGT	50
	(3) AG1-X2 (Dowex 1) Fractionation of Enzymatic Digests of Cyanogen Bromide Fragment Cn-1	52
	(4) Chromobead P Resin Fractionation of Basic, Neutral, and Acidic Peptides from Dowex 1 Column.	52

TABLE OF CONTENTS (cont'd)

CHAPTER			PAGE
II.	C.	High Voltage Paper Electrophoresis of Peptides	53
	D.	Amino Acid Analysis	55
		(1) Peptide Amino Acid Analysis . . .	55
		(2) Protein Amino Acid Analysis . . .	56
	E.	NH_2-Terminal and Sequence Determination of Peptides	56
		(1) Dansylation and Identification of Dansyl Amino Acids	57
		(2) Coupling with PITC (Phenylisothiocyanate)	62
		(3) Cleavage of PTC-peptide (Phenylthiocarbamylpeptide) . . .	63
		(4) Removal of Diphenylthiourea . . .	63
	F.	Disc-gel Electrophoresis.	64
	G.	Ultracentrifuge Studies	64
	H.	Optical Rotatory Dispersion and Circular Dichroism Measurements . . .	65
	I.	Enzyme Assays	68
		(1) Spectrophotometric Determination of BAEE Hydrolysis by SGT	68
		(2) The pH-stat Determination of ATEE and BAEE Hydrolysis by Pronase Enzymes	69
		(3) The pH-stat Determination of Oxidized Insulin A and B Chain Hydrolysis.	70
		(4) Spectrophotometric Determination of PNPA Hydrolysis by Pronase Enzymes	70

TABLE OF CONTENTS (cont'd)

CHAPTER		PAGE
II.	(5) Assay for Carboxypeptidase A Activity	71
	J. Digestive Procedures with α-Lytic Protease, Trypsin, α-Chymotrypsin and Pepsin	71
	(1) Digestion of Cyanogen Bromide Fragment Cn-1 from SGT with α-lytic Protease	71
	(2) Tryptic Digestion of Cn-1 Cyanogen Bromide Fragment. . . .	72
	(3) Tryptic Digestion of Performic Acid Oxidized SGT.	73
	(4) Chymotryptic Digestion of Cn-1 Cyanogen Bromide Fragment and Maleylated Performic Acid Oxidized SGT	74
	a. Digestion of Cn-1.	74
	b. Digestion of Maleylated Performic Acid Oxidized SGT.	74
	(5) Peptic Digestion of TLCK-SGT . .	75
	K. Chemical Modification of Proteins and Peptides	75
	(1) Cyanogen Bromide Cleavage of SGT	75
	(2) Reduction and S-β-Aminoethylation of Cyanogen Bromide Fragmentation Peptides from SGT	76
	(3) Maleylation of Amino Groups in Performic Acid Oxidized SGT and S-Aminoethylated Cn-1 Cyanogen Bromide Fragment.	77
	(4) Demaleylation of Amino Groups of Samples in Solution and on Paper	78
	(5) Acetylation of SGT.	78

TABLE OF CONTENTS (cont'd)

CHAPTER		PAGE
II.	(6) Carbamylation of the NH_2-Terminus of SGT with Potassium Cyanate. . .	79
	(7) Performic Acid Oxidation of SGT and Peptic Peptides of SGT on Paper	80
	(8) Active Site Directed Inhibition of SGT by TLCK.	81
	L. Diagonal Procedure for Identification of Disulphide Bond Peptides, ...	82
	M. Specificity Experiments with Oxidized Insulin A and B Chains.	82
	N. Automated Sequence Determination of SGT	84
	O. Determination of the Molar Extinction Coefficient of SGT.	85
	P. Stability Studies with SGT in the Presence and Absence of Ca^{2+}.	86
III.	RESULTS	88
	1. PHYSICO-CHEMICAL RESULTS	88
	A. Preparative Work	88
	(1) SE-Sephadex C-50 Chromatography of Pronase	89
	(2) CM-Sephadex C-50 Chromatography of Pronase	94
	B. Homogeneity of SGT	100
	C. Calculation of Molar Extinction Coefficient	103
	D. The Stability of SGT in the Presence and Absence of Ca^{2+}	103
	E. Determination of the Michaelis Constant for SGT.	109
	F. Enzyme Activity Versus pH Profile for SGT	112

TABLE OF CONTENTS (cont'd)

CHAPTER			PAGE
III.	G.	Optical Rotatory Dispersion and Circular Dichroism Studies.	115
	H.	Attempts to Block the NH_2-Terminal α-Amino Group by Acetylation and Carbamylation.	120
	I.	Molecular Weight Determination by Low Speed Sedimentation Equilibrium	124
	J.	Studies Relating to the Substrate Specificity of SGT.	125
	K.	Inhibition Studies with TLCK and PAB—Isolation of a Modified Histidine-57 Peptide.	138
	L.	Kinetic Isotope Effect of Deuterium Oxide on SGT Hydrolysis of BAEE	148
	2.	STRUCTURAL STUDIES	151
	A.	Introduction	151
	B.	Purification of Cyanogen Bromide Fragment Cn-1	156
	C.	Tryptic Digestion of Cyanogen Bromide Fragment Cn-1	168
	D.	α-Lytic Protease Digest of Cn-1 Fragment and Purification of Resultant Peptides	173
	E.	Automated 'Edman-Begg' Sequence Analysis of S-β-Aminoethylated Cyanogen Bromide Fragment Cn-1	191
	F.	Automated 'Edman-Begg' Sequence Analysis of Native SGT	200
	G.	Tryptic Digest of Performic Acid Oxidized SGT	205
	H.	Chymotryptic Digest of Maleylated S-β-Aminoethylated Cyanogen Bromide Fragment Cn-1	210
	I.	Chymotryptic Digest of Maleylated Performic Acid Oxidized SGT	218

TABLE OF CONTENTS (cont'd)

CHAPTER		PAGE
IV.	DISCUSSION	222
	BIBLIOGRAPHY	260

LIST OF TABLES

TABLE		PAGE
I.	Sequences Around Reactive Serine Residues. .	11
II.	Sequences in the Region of Serine-195 in Various Proteases	10
III.	Apparent Michaelis Constants for Various Trypsins and Trypsin-like Enzymes.	15
IV.	Physical Parameters of Various Mammalian Trypsins	17
V.	Sequences in the Region of Histidine-57 in Various Proteases	28
VI.	Sequences in the Region of Aspartic Acid-102 in Various Proteases	30
VII.	Elution System for Technicon Peptide Analyzer	51
VIII.	Moffit-Yang Parameters and α-Helix Estimations as a Function of pH.	114
IX.	Molecular Weight Determination by Sedimentation Equilibrium— Variation with Protein Concentration	125
X.	Major Peptides Resultant from Digestion of Oxidized Insulin A Chain (1:20 Molar Ratio).	132
XI.	Major Peptides Resultant from Digestion of Oxidized Insulin B Chain (1:20 Molar Ratio).	131
XII.	Positions of Hydrolysis of Oxidized Insulin A Chain by SGT Contaminant (1:20 Molar Ratio)	130
XIII.	Major Peptides Resultant from Digestion of Oxidized Insulin A and B Chains (1:1000 Molar Ratio)	129
XIV.	Position of 'Non-Tryptic' Hydrolysis of Oxidized Insulin A Chain by SGT Contaminant (1:1000 Molar Ratio)	128
XV.	Variation in Histidine Values after Inhibition with TLCK	145

LIST OF TABLES (cont'd)

TABLE		PAGE
XVI.	Partial Sequence of SGT Showing Missing or Ambiguous Regions	154
XVII.	Qualitative Solubility Tests on the Cyanogen Bromide Fragment Cn-1	161
XVIII.	Amino Acid Composition and Sequence of α-Lytic Peptides from Cyanogen Bromide Fragment Cn-1	184
XIX.	Summary of SGT Sequence and Alignment of α-Lytic Peptides from Cyanogen Bromide Fragment Cn-1.	175
XX.	Assumed Primary Sequence of SGT.	190
XXI.	Results of Automated Sequence Analysis of Cyanogen Bromide Fragment Cn-1.	194
XXII.	Results of Automated Sequence Analysis of Native SGT	203
XXIII.	Amino Acid Composition and Sequence of Chymotryptic Peptides from Cyanogen Bromide Fragment Cn-1	214
XXIV.	Homologous Sequences Adjacent to the NH_2-Terminus of SGT	231
XXV.	Homologous Sequences in the Region of the Disulphide Bridge Residues Half Cystine 42 and 58.	238
XXVI.	Homologous Sequences in the Region of Aspartic Acid-102.	236
XXVII.	Homologous Sequences in the Region of the Disulphide Bridge Residues Half Cystine 168 and 182.	234
XXVIII.	Homologous Sequences in the Region of the Active Serine-195.	244
XXIX.	Homologous Sequences in the Region of the Specificity Pocket	243
XXX.	Homologous Sequences in the Region of the COOH-Terminus of SGT	241

LIST OF TABLES (cont'd)

TABLE		PAGE
XXXI.	Comparison of Amino Acid Sequences of α-Lytic Protease, Porcine Elastase, Bovine Chymotrypsins A and B, Bovine Trypsin, and Streptomyces griseus Trypsin	252

LIST OF FIGURES

FIGURE		PAGE
1	Schematic mechanism of action of serine proteases	13
2	Position of dansyl amino acids on thin layer plates (Systems 1 and 2)	60
3	Position of dansyl amino acids on thin layer plates (Systems 1,2 and 3)	60
4	Position of dansyl dipeptides on thin layer plates (Systems 1 and 2)	58
5	Position of dansyl dipeptides on thin layer plates (Systems 1,2 and 3)	58
6	Bio-Rex-70 chromatography of CM-cellulose fractionated Pronase	87
7	SE-Sephadex chromatography of Pronase	90
8	CM-Sephadex chromatography of Pronase	92
9	SE-Sephadex chromatography of CM-Sephadex purified SGT	97
10	Bio-Rex-70 chromatography of CM- and SE-Sephadex purified SGT	96
11	Polyacrylamide gel electrophoresis of SGT	99
12	Stability of SGT activity with pH at 0° in the absence of calcium ion	102
13	Stability of SGT activity with pH after 25 min at 30° in the presence and absence of calcium ion	105
14	Stability of SGT activity with pH after 24 hrs at 30° in the presence and absence of calcium ion	104
15	Lineweaver-Burk plot for estimation of the SGT apparent Michaelis constant	108
16	Plot of the pH dependence of bovine trypsin activity	111

LIST OF FIGURES (cont'd)

FIGURE		PAGE
17	Plot of the pH dependence of SGT activity	110
18	Plot of a_o versus pH for SGT.	113
19	Circular dichroism spectra of SGT observed over the pH range 5.3 to 7.6.	118
20	Circular dichroism spectra of SGT observed over the pH range 8.5 to 10.5	117
21	Optical rotatory dispersion spectra of SGT observed over the pH range 7.3 to 10.1.	116
22	Low speed sedimentation equilibrium ln y versus r^2 plot for determination of molecular weight . .	123
23	Bio-Rex-70 chromatography of SE-Sephadex purified SGT.	126
24	Electropherogram of SGT and TLCK inhibited SGT digests of oxidized insulin B chain	135
25	Electropherogram of SGT and TLCK inhibited SGT digests of oxidized insulin A chain	134
26	Time course of inhibition of SGT with TLCK. . . .	140
27	Competitive inhibition of TLCK inhibition of SGT by p-aminobenzamidine	139
28	A pH 6.5/ pH 6.5 diagonal peptide map of a peptic digest of TLCK inhibited SGT	143
29	A pH 6.5/ pH 6.5 diagonal peptide map of a peptic digest of SGT.	142
30	Plots of the pH and pD dependence of SGT activity.	149
31	Sephadex G-50 gel-filtration of cyanogen bromide cleavage products from SGT	152
32	Sephadex G-75 Superfine gel-filtration of cyanogen bromide cleavage products from SGT . . .	157
33	Sephadex G-75 Superfine gel filtration of cyanogen bromide cleaved SGT previously purified on CM- and SE-Sephadex	160

LIST OF FIGURES (cont'd)

FIGURE		PAGE
34	Sephadex G-75 Superfine gel-filtration of cyanogen bromide cleavage products from SGT in 0.65 M acetic acid/8 M urea	165
35	Sephadex G-75 Superfine gel-filtration of cyanogen bromide cleavage products from SGT in 0.65 M acetic acid/8 M urea (preparative run)	167
36	SE-Sephadex ion exchange chromatography of insoluble pellet from tryptic digest of Cn-1 fragment	170
37	Variation in turbidity with time during digestion of Cn-1 fragment with α-lytic protease	172
38	Chromatography of fraction I from the α-lytic digest of Cn-1 fragment on peptide analyzer	187
39	Chromatography of fraction II from the α-lytic digest of Cn-1 fragment on peptide analyzer	186
40	Chromatography of fraction III from the α-lytic digest of Cn-1 fragment on peptide analyzer	185
41	Composite plot of yields for six PTH-amino acids obtained during the automated sequential degradation of Cn-1 fragment (Residues 48 to 59)	192
42	Composite plot of yields for seven PTH-amino acids obtained during the automated sequential degradation of SGT (Residues 41 to 60)	201
43	Sephadex G-50 gel-filtration of a tryptic digest of performic acid oxidized SGT	206
44	Chromatography of fraction I from the chymotryptic digest of Cn-1 fragment on peptide analyzer	216
45	Chromatography of fraction III from the chymotryptic digest of Cn-1 fragment on peptide analyzer	215
46	A pH 6.5/ pH 6.5 lysine diagonal peptide map of a chymotryptic digest of performic acid oxidized SGT	219

LIST OF FIGURES (cont'd)

FIGURE		PAGE
47	Ribbon diagram of chymotrypsin showing homologous residues in the region of the NH_2-terminus of SGT	230
48	Ribbon diagram of chymotrypsin showing homologous residues in the region of half cystine residues 42 and 58 of SGT	237
49	Ribbon diagram of chymotrypsin showing homologous residues in the region of aspartic acid-102 of SGT.	235
50	Ribbon diagram of chymotrypsin showing homologous residues in the region of half cystine residues 168 and 182 of SGT	233
51	Ribbon diagram of chymotrypsin showing homologous residues in the region of serine-195 and the specificity pocket of SGT.	242
52	Ribbon diagram of chymotrypsin showing homologous residues in the region of the COOH-terminus of SGT	240
53	Speculative diagram of SGT compared with a schematic α-chymotrypsin molecule	246

LIST OF ABBREVIATIONS

AECys	S-β-aminoethylcysteine
ATEE	N-acetyl-L-tyrosine ethyl ester
BAEE	N-benzoyl-L-arginine ethyl ester
CD	Circular dichroism
CM	Carboxymethyl-
CM-His	Carboxymethylhistidine
DFP	Diisopropylfluorophosphate
DIP	Diisopropylphosphoryl-
DNS-Cl	1-dimethylaminonaphthalene-5-sulphonyl chloride
DNS	1-dimethylaminonaphthalene-5-sulphonyl
D_2O	Deuterium oxide
$Met(O_2)$	Methionine sulphone
MW	Weight-average molecular weight
ORD	Optical rotatory dispersion
PAB	p-aminobenzamidine
PITC	Phenylisothiocyanate
PTC-peptide	Phenylthiocarbamyl-peptide
SA	Specific activity
SE	Sulphoethyl-
TFA	Trifluoroacetic acid
TLCK	L-(1-tosylamido-2-lysyl) ethylchloromethyl ketone
TPCK	L-(1-tosylamido-2-phenyl) ethylchloromethyl ketone

All temperatures are given in degrees Centigrade

CHAPTER I

INTRODUCTION

Proteolytic enzymes are found in both eukaryotic and prokaryotic organisms and are responsible for the hydrolysis of peptide bonds during the degradation of proteins and peptides. A classification of such a ubiquitous group of enzymes can be made by first examining those which are present in vertebrates, followed by those present in plants and lastly those found in microbial organisms.

Proteases from any of the above classes of organisms can be further subdivided into two groups; those which hydrolyze only the NH_2-terminal or COOH-terminal residues in a polypeptide chain, termed exopeptidases, and those which hydrolyze internal peptide bonds, termed endopeptidases. However, a more useful classification scheme for discussion of the vertebrate proteases would be one based on those enzymes associated with the digestive processes and one based on those enzymes external to this system. Examples of the latter group would include the cathepsins, kallikreins and kininases, while the former would include the well-known pancreatic and gastrointestinal enzymes. It is this latter group of enzymes which I would like to discuss initially.

The pancreatic proteases have probably been studied more than any other single group of enzymes. This is

partly because they were readily available in a concentrated package which facilitated purification, and partly because of the interest investigators had in utilizing them for their inherent catalytic properties. As stated previously, proteolytic enzymes can be divided into 'exo' and endopeptidases. This is also true of enzymes in the pancreas. In this organ, inactive precursor molecules called zymogens are synthesized and stored in intracellular organelles. The pancreatic exopeptidases are represented by the zymogens procarboxypeptidase A and procarboxypeptidase B. Both carboxypeptidases are zinc-containing proteolytic enzymes which remove the COOH-terminal amino acid from peptide chains. The A enzyme prefers aromatic and long chain aliphatic residues while the B enzyme hydrolyzes lysine and arginine most rapidly (1). The pancreatic endopeptidases are represented by trypsinogen, chymotrypsinogens A and B as well as proelastase from pig pancreas (2). Activation of these inactive precursors is usually catalyzed by trypsin. Trypsinogen, however, is activated by both trypsin and enterokinase (1). Chymotrypsin A and B have almost identical enzymatic specificities, while, on the other hand, there are large differences in the specificities of the chymotrypsins, trypsin and elastase (1). These and other functional aspects relevant to this group of proteases will be discussed in depth at a later time. It is important,

however, to indicate that these enzymes belong to a class of proteases termed serine proteases, since in each enzyme a reactive serine-195 (α-chymotrypsin numbering system) reacts with organophosphorous compounds, such as diisopropylfluorophosphate (DFP), inhibiting the enzyme (1). The same serine residue is acylated as a step in the hydrolysis of ester substrates (3). Hydrolysis of substrates by all pancreatic serine proteases and most serine proteases in general, is optimally carried out at alkaline pH values.

Still other proteolytic enzymes are products of the gastrointestinal tract and can be easily distinguished on this basis from the enzymes produced in the pancreas. This group of enzymes can be divided into those which originate in the stomach and those which are elaborated in the mucosa of the intestine. The latter enzymes are typified by leucine aminopeptidase and enterokinase. Leucine aminopeptidase is the classical NH_2-terminal exopeptidase, first observed in swine intestinal mucosa extracts, by Linderstrøm-Lang in 1929 (4). Although it will remove many different NH_2-terminal residues, leucine is removed very rapidly and thus the name leucine aminopeptidase is retained to distinguish it from other aminopeptidases.

Enterokinase, on the other hand, is an intestinal endopeptidase which has some very distinctive characteristics. As obtained from swine duodenal extract, the

enzyme has been shown to be a glycoprotein which catalyzes the activation of trypsinogen much more efficiently than trypsin (5). Maroux et al (1971) have also shown that the K_m for enterokinase, using bovine trypsinogen as a substrate, is six times lower than the K_m for trypsin (5). The former investigators also demonstrated that the rate constant for the rate determining step (k_{cat}) was 2000 times greater for enterokinase than for trypsin. Apparently, enterokinase is similar to trypsin in that it is inhibited by both DFP and N-tosyl-L-lysine chloromethyl ketone (TLCK), suggesting that like the serine proteases, an active serine and a histidine residue participate in the catalytic reaction. On the other hand, these authors claim that in contrast with trypsin, the sequence after the NH_2-terminal valine, -Asp-Asp-Asp-Asp-Lys-, appears to be essential for the interaction of enterokinase with peptide or protein substrates, resulting in the high degree of specificity exhibited by this enzyme. For example, it is known that chymotrypsinogen A is not activated and the peptide Val-Ala-Ala-Lys-Ile-Val-Gly is not hydrolyzed (5). Enterokinase cleaves exclusively the bond between residues 6 and 7 (-Lys-Ile-) in trypsinogen.

The other gastrointestinal enzymes to be recognized, are those found in the stomach. They are active at acid pH values (pH 1-5) and are therefore, termed collectively, acid proteinases. These enzymes, also termed pepsins, are elaborated by specific cells in the stomach mucosa and

may be a group of very similar proteases, which have apparent differences arising out of varying degrees of proteolysis and phosphorylation of the activated zymogen(s) (6). In addition to this principal enzymic component, a 'pepsin-like' enzyme named gastricin is also present in the stomach. However, it is not clear whether or not this is a deserved differentiation, as certain evidence tends to indicate that a human pepsin 'I' may be identical with gastricin (6). This group of proteases is inhibited by α-diazo-p-bromoacetophenone where 1 mole of inhibitor is bound per mole of enzyme (7).

The preceding material has dealt with those vertebrate proteases associated with the gastrointestinal tract. Earlier, another group of vertebrate proteases was alluded to, and it is this group which I wish to quickly review at this time. This class of enzymes can be divided into the cathepsins, kallikreins and kininases. The cathepsins represent a complex group of lysosomal enzymes which are derived from a variety of animal tissue extracts (8). While there is still much unknown about the cathepsins, it is known that as a group, they contain both 'endo' and exopeptidases. Cathepsins A and C are exopeptidases, the former being a carboxypeptidase and the latter an aminopeptidase, while cathepsins B,D, and E appear to have endopeptidase activity (9). All of the cathepsins have been isolated from the spleen, although some, such as the

sulphydryl enzyme cathepsin B, are found in many tissues (9).

The kallikreins are the second group of proteases whose action is exclusive of the digestive process. They are present in activatable forms in the vascular system and throughout the glandular organs of the body (9). Most kallikreins show little activity towards casein, but hydrolyze the ethyl and methyl esters of benzoyl-L-arginine and the methyl ester of tosyl-L-arginine (9). When activated by either Hageman's factor or plasmin, the plasma kallikreins increase the peripheral concentration of hypotensive vasoactive kinins, from α-2-globulins, by a proteolytic mechanism (9,10). Kallikrein(s) are inhibited by DFP and soybean trypsin inhibitor, indicating that they may be related to the serine proteases——with specific reference to trypsin (10).

The products of kallikrein proteolysis are active in nanogram quantities and so it is not surprising that still another group of plasma and tissue enzymes destroy these peptides by a carboxypeptidase activity (9). These kininases are present in white cells, spleen, lung and free in the plasma. The lung appears to be the main source of kinin inactivation (9).

As indicated initially, certain plant proteases have also been examined in great detail. Examples of both sulphydryl and serine proteases can be clearly demonstrated. The former group is typified by papain which is isolated from papaya latex and requires a free

sulphydryl for activity (11). Other sulphydryl proteases of plant origin are chymopapain, also isolated from papaya latex, ficin from fig trees and stem bromelain from pineapple (11). There may also be some relationship between certain sulphydryl cathepsins in animal tissues with these plant enzymes, as all of these proteases require an activator such as cyanide, cysteine, or glutathione in order to release the blocked thiol group required for activity (11).

Examples of DFP-sensitive proteases isolated from plants, include two serine proteases extracted from French beans. Caseinase and phaseolin have the following sequences around the reactive serine residue: -Thr-Ser-Met-Ala- for caseinase and -Glu-Ser-Val- for phaseolin. These sequences are similar to the active site sequences found in the subtilisins, with respect to caseinase, and Baker's yeast, with respect to phaseolin (12).

As indicated in the introduction to this section, serine proteases have been isolated from microbial sources as well as vertebrate and plant sources. Indeed, together with the sulphydryl proteases, the metallo-proteases, and acid proteases, the microbial world is a rich source of proteolytic enzymes. DFP-sensitive endopeptidases alone, have been isolated from Bacillus, Arthrobacter, Streptomyces, Aspergillus, Penicillium, Bacteroides, Sorangium, Alternaria, and Saccharomyces organisms (13). It is

important at this time, to indicate that the subtilisins, obtained from several Bacillus species, have been shown to have a reactive serine inhibited by DFP and give every indication of having a catalytic mechanism very similar to the Asp-Ser-Gly serine proteases (14). The subtilisins have an active serine sequence of Thr-Ser-Met-Ala and it is known that a histidine residue participates in the catalytic mechanism——again, like the Asp-Ser-Gly serine proteases (14). However, the subtilisins appear to have no similarity in amino acid sequence with the Asp-Ser-Gly group of proteases and appear to have evolved independently in a convergent fashion, to finally share a similar mechanism of action.

Sulphydryl proteases have been isolated most notably from Streptococcus and Clostridium species, while acid proteases similar to pepsin appear to be most notably found amongst the molds. The latter group of proteases are not inhibited by DFP, thiol poisons or metal chelators (13). The metallo-endopeptidases as a group, have been isolated from Bacillus, Streptomyces, Aspergillus, Pseudomonas, Proteus, Micrococcus and many other microbial genera. All appear to share a common specificity demonstrated by hydrolysis on the NH_2-terminal side of residues with hydrophobic side chains. In addition, all appear to be inhibited by metal chelators while being insensitive to either DFP or thiol reagents (13).

Thermolysin from Bacillus thermoproteolyticus is a notable example of this group whose primary sequence has been recently ascertained by Titani et al (1972) (15).

Since the enzyme examined in this dissertation, Streptomyces griseus trypsin (SGT), has been found to be an Asp-Ser-Gly serine protease inhibited by DFP, a more thorough discussion of this class of proteases is required (16). The source of this enzyme is Streptomyces griseus pronase, a commercial extracellular filtrate, composed of a complex mixture of 'exo' and endopeptidases together with material lacking proteolytic activity (17). Although the carboxypeptidase and aminopeptidase activities have not been studied in detail (18,19,20), three of the endopeptidases have been shown to be Asp-Ser-Gly serine proteases (16,21).

Various methods of separation have been employed in an attempt to purify the principals of pronase, but due to its apparent complexity none have been, nor probably will be, entirely successful. The CM-Sephadex system devised by this laboratory seems to resolve the largest variety of constituents in the most efficient manner to date (20). The development of this preparative column procedure, utilizing a volatile pyridinium acetate buffer system, will be discussed in this dissertation. One of the products of this system, and the enzyme the system was primarily designed to resolve, is Streptomyces griseus

TABLE II

SERINE-195 SEQUENCES

Protease	Sequence (195)
Trypsin-Bovine	Asp-Ser-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Val-Val-Cys
Trypsin-Porcine	Asp-Ser-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Val-Val-Cys
Trypsin-Shark (Spiny Dogfish)	Asp-Ser-Cys(Glx,Gly,Asx,Ser,Gly,Gly,Pro,Val)Val-Cys
Chymotrypsin A-Bovine	Ser-Ser-Cys-Met-Gly-Asp-Ser-Gly-Gly-Pro-Leu-Val-Cys
Chymotrypsin B-Bovine	Ser-Ser-Cys-Met-Gly-Asp-Ser-Gly-Gly-Pro-Leu-Val-Cys
Elastase-Porcine	Ser-Gly-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Leu-His-Cys
Thrombin-Bovine	Asp-Ala-Cys-Glu-Gly-Asp-Ser-Gly-Gly-Pro-Phe-Val-Met
Plasmin-Human	Ser-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Leu-Val-Cys
SGT[1]	Asp-Thr-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Met-Phe-Arg
SGPA[2]	Cys-Ala-Gln-Pro-Gly-Asp-Ser-Gly-Gly-Ser-Leu-Phe-Ala
α-Lytic Protease-Myxobacter 495	Cys-Met-Gly-Arg-Gly-Asp-Ser-Gly-Gly-Ser-Trp-Ile-Thr

[1] *Streptomyces griseus* trypsin
[2] *Streptomyces griseus* protease A (29)

TABLE I

SEQUENCES AROUND REACTIVE SERINE RESIDUES

Source	Sequence
Various Mammalian Enzymes	Asp-Ser-Gly
Sorangium sp.	Asp-Ser-Gly
Streptomyces griseus	Asp-Ser-Gly
Subtilisins	Thr-Ser-Met-Ala
Caseinase (French bean)	Thr-Ser-Met-Ala
Aspergillus	Thr-Ser-Met-Ala
Baker's yeast	Glu-Ser-Val
Phaseolin (French bean)	Glu-Ser-Val

trypsin, which in the following pages, will be shown to have kinetic and physical properties not unlike bovine trypsin and other Asp-Ser-Gly proteases.

As a group, the serine proteases give every indication of being similar in their basic catalytic mechanism. This is true in spite of slight differences observed in those areas of the primary sequence about the active serine. Table I indicates some of these differences as represented by Markland and Smith (1971) (12). Table II shows a collection of sequences from various enzymes belonging to the Asp-Ser-Gly group of proteases and associated with the region around the active serine 195 (22). Some of the SGT sequence has been extended beyond that published in the cited reference. Sequences about other active site residues, to be presented later, have been similarly extended. (All sequence numbering is that of bovine α-chymotrypsin).

As indicated earlier, the active site serine in the serine proteases is transiently acylated by the carbonyl moiety of the peptide bond being cleaved (23,24). The kinetic mechanism for the serine proteases has been intensively studied and the basic kinetic scheme, primarily derived from experiments with α-chymotrypsin, is as shown below (23,25,26,27,28).

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES' \xrightarrow{k_2} ES'' \xrightarrow{k_3} E + P_2$$

$$ES'' \downarrow P_1$$

FIGURE 1

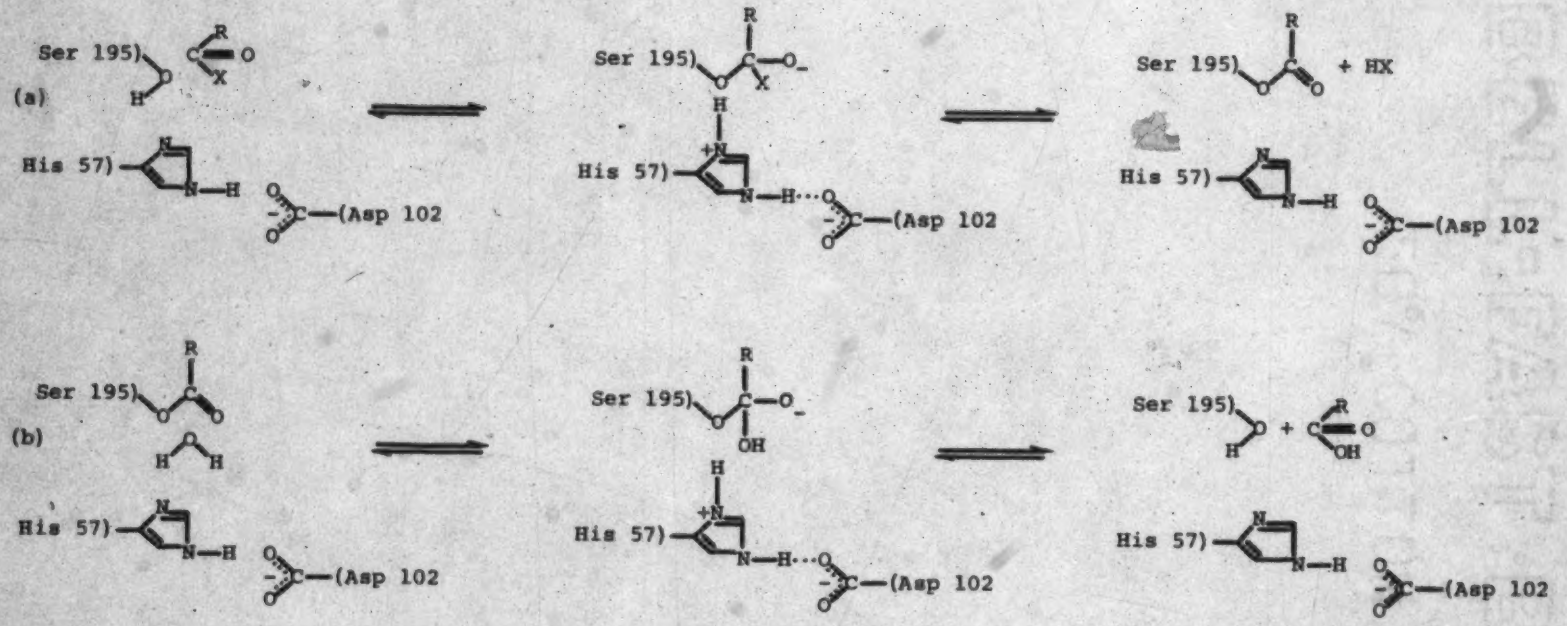


Fig. 1: Possible mechanism of the formation (a) and the hydrolysis (b) of the acyl enzyme in the catalysis by serine proteases; X represents the leaving group. From Polgar and Bender (32).

In this scheme, ES is the enzyme-substrate complex, ES' the acyl-enzyme intermediate, P_1 the leaving group of the substrate or product number one, and P_2 the carboxylic acid (27). The acylation step (k_2) is a nucleophilic reaction dependent on two groups, one with a pk_a of 7 and the other with a pk_a of 9 or 10, while the deacylation step (k_3) is dependent on only one group with a pk_a of 7 (30). This latter group has been found to be one of the histidines in trypsin and α-chymotrypsin by modification studies which will be elaborated upon later (31). In a similar manner, the active serine-195, acylated in the reaction mechanism shown above, was found to be specifically phosphorylated by DFP (1). This reaction irreversibly inactivates all serine proteases.

The rate constant parameters for the reaction mechanism previously described, vary depending upon whether an ester or an amide is used as a substrate (25). When S is an amide, k_2 or the acylation step is rate limiting and P_1 is an amine; however, when S is an ester, as is most often the case in this thesis, k_3 or the deacylation step is rate limiting and P_2 is an alcohol. Figure 1 is a schematic representation of the possible charge relay mechanism of formation (a), and hydrolysis (b), of the acyl-enzyme in the catalysis by a generalized serine protease as proposed by Polgar and Bender (1969) (32). The mechanism in Figure 1 incorporates features

TABLE III

APPARENT MICHAELIS CONSTANTS FOR VARIOUS TRYPSINS AND TRYPSIN-LIKE ENZYMES

Enzyme	Substrate*			Reference No.
	BAEE Km(mM)	TAME Km(mM)	BAA Km(mM)	
Streptomyces fradiae Trypsin	0.0015	—	0.32	37,38
Bovine Trypsin	0.0043	0.0125	3.1	39,40,41
Bovine Thrombin	—	0.0016	—	42
Porcine Kallikrein	—	0.006	—	43
Stem Bromelain	0.17	—	0.0012	44
Streptomyces erythreus Trypsin	0.0018	—	—	37

* BAEE (N-Benzoyl-L-Arginine Ethyl Ester)

TAME (N-Tosyl-L-Arginine Methyl Ester)

BAA (N-Benzoyl-L-Argininamide)

of those proposed by Wang (1968) (33) and Blow et al (1969) (24) and is completely consistent with the geometry of the residues indicated, as deduced by X-ray diffraction analysis of the crystalline α-chymotrypsin enzyme (24). It is also consistent with the geometry established from the tertiary structure determinations of porcine elastase (34) and bovine trypsin (35). The results of these tertiary structure determinations will be examined in depth at a later time, but they are compatible with the hypothesis that all Asp-Ser-Gly serine proteases have the same basic catalytic mechanism.

In connection with physico-chemical studies undertaken in this thesis, it is important to briefly indicate results obtained from other laboratories investigating serine proteases. For example, a collection of apparent Michaelis constants for various trypsins and 'trypsin-like' enzymes is shown in Table III while Table IV lists some molecular weights and isoionic points determined for three mammalian trypsins. It has been shown that the isoionic or isoelectric points of all trypsins do not fall within the alkaline pH range. For example, rat, dogfish, and lungfish trypsins are anionic at neutral pH, whereas bovine and porcine trypsins are cationic as indicated in Table IV (36). Thus the net charge on trypsin molecules does not appear to be important for the specific function of these enzymes.

TABLE IV

PHYSICAL PARAMETERS OF MAMMALIAN TRYPSINS

Enzyme	Apparent Molecular Weight	Isoionic Point	Reference No.
Bovine Trypsin	24,000	10.1	48,49
Porcine Trypsin	23,400	10.8	50
Human Trypsin	22,900	—	51

Various specificity and stability studies will be discussed in the results to follow, which attempt to indicate the similarity of SGT with bovine trypsin and other serine proteases. One such experiment involved the role of Ca^{2+} ion in stabilizing the enzyme towards autolysis. It has long been known that calcium stabilizes trypsin from autolysis (45). Recently, Sipos and Merkel (1970) have proposed from the development of a positive differential ultraviolet spectrum in the prescence of calcium ion and from optical rotatory dispersion (ORD) studies that the function of calcium ion is to maintain a specific compact conformation of the molecule (46). This idea had previously been expounded by Lazdunski and Delaage (1967) from spectrophotometric titration data (47). The proposal has also been corroborated in principle by the X-ray diffraction work of Stroud et al (1971) who indicate that the geometry of the

trypsin molecule will allow calcium ions to form ionic linkages between pairs of anionic groups within a specific loop or between a loop and the surface of the molecule, tending to make a more rigid, less accessible structure (35).

The activation of zymogens is affected by the presence of calcium ions and the discussion of zymogen activation, which follows, necessarily must include this phenomenon (52). For example, in the process of trypsinogen activation, as much as 50% of the potential enzyme activity is lost as enzymatically inactive products, in the absence of calcium (53). The rate of selective hydrolysis of the Lys 6-Ile 7 bond by trypsin is increased by 20 mM Ca^{2+}, while on the other hand, the amount of nonspecific cleavage is decreased (53). It is worth recalling, at this time, that although it has been shown by Maroux et al (1971) that the four adjacent aspartyl residues have a negative effect on the activation of trypsinogen by trypsin, the activation of this zymogen by enterokinase depends upon this sequence (5). Abita et al (1969) have indicated that calcium ion may mitigate this negative effect of the cluster of aspartic acids by binding to the four aspartyl residues (54).

During the autocatalytic activation of trypsinogen in alkaline media a family of products is formed. The primary product is β-trypsin resulting from the hydrolysis of the previously mentioned bond between lysine-6 and

isoleucine-7 releasing a hexapeptide (53). Various investigators have shown that several other bonds can be cleaved, resulting in active enzymes. For example, hydrolysis of β-trypsin at the Lys 131-Ser 132 bond leads to α-trypsin, while subsequent cleavage of the Lys 176-Asp 177 bond yields pseudotrypsin, another active form of trypsin (53).

Chymotrypsinogen is similarly activated by hydrolysis of the Arg 15-Ile 16 bond by trypsin to form the fully active π-chymotrypsin (55). After this tryptic activation, all subsequent products result from autodegradation of first, π-chymotrypsin, then δ-chymotrypsin, finally producing γ or α-chymotrypsin. No evidence for a structural or enzymatic difference between the γ and α forms of the enzyme exists. An inactive neo-chymotrypsinogen formed by autodigestion of chymotrypsinogen can be activated by incubation with trypsin to also form α-chymotrypsin (55). It should be mentioned at this time, that Oppenheimer et al (1966) have given evidence to indicate that the isoleucine-16 NH_2-terminal α-amino group of chymotrypsin is required to maintain the enzyme in an active conformation (56). This result was corroborated by Sigler et al (1968), when the X-ray diffraction determination of the tertiary structure of bovine α-chymotrypsin clearly showed that the free α-amino group of isoleucine-16 formed an internal ion pair with aspartic acid-194 in the active site (57).

Stroud et al (1971) have also shown that the NH_2-terminal α-amino group of bovine trypsin forms a similar ion pair with aspartic acid-194 in this molecule (35). The function of the NH_2-terminal α-amino group, is not clearly defined in all the serine proteases, however, and will be reviewed in depth in the following discussion.

As mentioned earlier, a group of residues in the serine proteases has been implicated in the catalytic mechanism. They have been termed the active site residues and include serine-195, aspartic acid-102, histidine-57, aspartic acid-194, and quite possibly the NH_2-terminal residue. Other residues close to this active site form the 'specificity pocket' and are required by the enzyme to effect binding of specific substrates. I would like to discuss all these important residues in order, beginning with aspartic acid-189 found in the specificity site of trypsin and continuing on to discuss the active site residues.

Bovine trypsin has been in the shadow of α-chymotrypsin for a considerable length of time and only in the last two years has there been X-ray data available for this molecule (35). Although the mechanism of action of this serine protease appears basically the same as α-chymotrypsin, its substrate specificity is markedly different (58). Unlike the latter enzyme, which prefers to hydrolyze peptide bonds on the COOH-terminal side of aromatic residues and to a lesser extent, aliphatic

residues, bovine trypsin is much more specific. Hydrolysis occurs only at peptide bonds on the COOH-terminal side of arginine and lysine residues (59). The source of this specificity has been associated with a charged aspartic acid residue buried in the active site, as inferred by experiments conducted by Eyl and Inagami (1970) (60) and corroborated by Stroud and co-workers during their X-ray analysis of crystalline DFP inhibited bovine trypsin (35). No such charged residue exists in this position in α-chymotrypsin; instead it is replaced by a serine residue (57). Stroud et al (1971) have stated, that from model-building experiments with lysl or arginyl amides, it is evident that aspartic acid-189 is ideally situated to form a charge interaction with the substrate side chain at the back of the binding pocket (35). The restraint of this charge interaction, together with a hydrogen bond between the carbonyl of serine-214 and a proton of the last N-H group on the substrate, bring the bond to be hydrolyzed close to both the nucleophilic oxygen of serine-195 and the histidine-57 imidazole (35). The diffraction studies also indicated differences in the architecture of the specific binding sites of trypsin and α-chymotrypsin, which were alike, and elastase which was different. The latter protease has bulky substitutions in the region adjacent to the binding site, which suppress binding of large side chains (35). This is in agreement with the difference in substrate side

chain specificities observed with these enzymes. Further observations regarding still other residues associated with the specificity site in crystalline DIP-trypsin will be discussed when the tertiary structure of bovine trypsin is compared with α-chymotrypsin, later in this chapter.

Serine-195, or the active serine, was the first residue in the active site to be distinguished as fundamental to the activity of the serine proteases. As early as 1949, Jansen et al reacted DFP with chymotrypsin to form the inactive diisopropylphosphoryl-enzyme (61). Subsequent sequence studies indicated that a single serine residue (Ser 195) was phosphorylated (62). Since that time many other serine proteases have been identified and shown to be similarly sensitive to DFP (63). Phenylmethyl sulphonyl fluoride and several other reagents have also been used to specifically inactivate serine proteases by modification of serine-195 (64). Other evidence to suggest that this serine was involved in catalysis and extraordinarily reactive, came from kinetic studies where various acyl-enzyme intermediates were characterized. Both cinnamoyl-α-chymotrypsin and cinnamoyl-trypsin have been prepared by Bender and co-workers by following the reactions of N-trans-cinnamoylimidazole spectrophotometrically. It was found necessary to conduct the reaction at pH 5.2, where spontaneous hydrolysis and deacylation of the acyl-intermediate were minimized (65). Similarly,

deacylation of the acyl-intermediate formed on reaction of chymotrypsin with p-nitrophenyl acetate was shown to be rate limiting, by demonstrating that a burst of p-nitrophenol product, equivalent to the molar concentration of the enzyme, was produced on introduction of the substrate (66). This was followed by a slow steady state liberation of p-nitrophenol resulting from hydrolysis of the acyl-enzyme.

Since it was known from work done in Wählby's laboratory in Sweden that the 'trypsin-like' enzyme from Streptomyces griseus was an 'Asp-Ser-Gly' serine protease, inhibited by DFP, work on the remaining kinetically important residues in SGT was emphasized in this dissertation (16). Specifically, we were interested in the active site histidine and the NH_2-terminal α-amino group. With respect to the latter group, several laboratories have observed that with α-chymotrypsin, the state of ionization of the α-amino group appears to control the active conformation and substrate binding ability of the enzyme. For example, hydrolysis of the single Arg 15-Ile 16 bond in the activation of chymotrypsinogen by trypsin is associated with the appearance of enzyme activity (67). In addition, Hess and colleagues (1967), demonstrated that while the rate constant for the rate determining acylation of α-chymotrypsin, using a neutral amide as substrate, was independent of pH in the alkaline range, the $K_{m(app)}$ was found to be pH dependent (68).

Since according to these investigators, the $K_{m(app)}$ for hydrolysis of a specific amide substrate was a measure of an overall enzyme-substrate dissociation constant, it was concluded that an ionizing group with $pk_{(app)}$ of about 8.5 affected the enzyme-substrate dissociation constant, and not the bond breaking step. Oppenheimer et al (1966) have also shown that fully acetylated chymotrypsinogen can be activated to yield acetylated δ-chymotrypsin (56). Furthermore, the specific rotation of this acetylated enzyme was observed to vary with pH in the region 6.0 to 10.5, the change in specific rotation following the ionization of a group with $pk_{(app)}$ of 8.3. Subsequent reacetylation of the acetylated enzyme resulted in inactivation, apparently as a result of modification of the NH_2-terminus. McConn et al (1969) carried out a more extensive conformational study of acetylated chymotrypsinogen and active acetylated δ-chymotrypsin by optical rotatory dispersion (ORD) and circular dichroism (CD) measurements (69). These studies indicated that at high pH, where the enzyme is inactive and the α-amino group of isoleucine-16 is deprotonated, the enzyme assumes the same conformation as the zymogen. The preceding data indicate the essentiallity of a protonated α-NH_2 group for the activity of the enzyme and are consistent with its known structure in the crystalline state where X-ray diffraction analyses have shown that the charged α-NH_2 group of isoleucine-16 is in a buried position and forms

an internal ion pair with aspartic acid-194 (57). Similar ion pairs involving the NH_2-terminal residues of elastase (34) and trypsin (35) have also been observed in the X-ray diffraction analyses of these enzymes.

However, the essential nature of this ion pair for the activity of these enzymes may be questioned. Valenzuela and Bender (1969, 1970) have studied the δ-chymotrypsin-catalyzed hydrolysis of specific substrates and the binding properties of three competitive inhibitors with α- and δ-chymotrypsins over the pH range 7 to 11 (70, 71). Unlike α-chymotrypsin, δ-chymotrypsin was shown to be active and to bind substrates and inhibitors to a significant extent when isoleucine-16 is not protonated. For example, it was possible to show that δ-chymotrypsin at pH 11.5 binds N-acetyl-D-tryptophanamide with a K_s equal to 8×10^{-3} M. Agarwal, Martin, Blair, and Marini (1971) have come to the same conclusion based on studies of δ-chymotrypsin modified at the NH_2-terminal isoleucine-16 by reaction with either ethyl acetimidate or methyl picolinimidate (72). The resulting amidinated enzyme preparations showed no change in enzyme activity when measured at pH 8.0 with the specific substrate N-acetyl-tyrosine ethyl ester. Using N-acetyl-L-tryptophan ethyl ester as substrate, the amidinated enzyme showed a dependence on pH with $pK_{(app)}$'s equal to 6.7 and 9.3. They concluded that the protonated amino group of isoleucine-16 is not essential for chymotryptic activity

and is not the group with $pK_{(app)}$ of about 9 seen on acylation or on binding of substrate.

With elastase also, the importance of the protonated α-NH_2 group seems questionable. Although in the crystalline state there is no question that the NH_2-terminal residue forms an internal ion pair with aspartic acid-194, Kaplan and Dugas have shown that both native and fully acetylated elastase remain fully active at all pH values between 8 and 10 (73). This is in spite of the fact that the pK_a of the NH_2-terminus has been shown to have a pK_a of 9.7 (74). The low reactivity of this group however, even at pH 10.0, indicates that even when deprotonated it remains largely buried in the molecule until denatured at higher pH. The enzyme denatures at pH 10.5 (73).

Other serine proteases also appear to depart from the required NH_2-terminal participation seen in α-chymotrypsin. The α-lytic protease of Myxobacter 495, like elastase, shows no conformational change with pH between 5.0 and 10.5 (75) and both the native and acetylated enzyme are fully active between pH 8.0 and 10.0 (76).

In the case of bovine trypsin the experimental evidence concerning the importance of the protonated NH_2-terminal residue is more fragmentary. Stroud et al (1971), as mentioned earlier, have recently determined the tertiary structure of crystalline bovine DIP-trypsin

and shown that the NH_2-terminal α-amino group is buried in the active site and forms an ion pair with aspartic acid-194 (35). However, unlike α-chymotrypsin, trypsinogen is not amenable to activation after acetylation, and reaction at the NH_2-terminal α-amino group of trypsin is not possible with acetic anhydride (77). As a result of this problem, Scrimger and Hofmann (1967) utilized a different approach and demonstrated that deamination of the α-amino group with nitrous acid at pH 4.0 caused inactivation of the enzyme (78). More recently, Robinson et al (1973) demonstrated that an active guanidinated trypsin lost essentially all activity on carbamylation of the NH_2-terminal α-amino group using potassium cyanate, corroborating the observation of Hofmann's laboratory (79). However, no definitive study of the pH dependence in the alkaline range is known to have been done with bovine trypsin. This is presumably due to its relative instability at high pH.

The histidine-57 in trypsin has been identified with a pH dependent residue which has a $pK_{(app)}$ of approximately 7.0 on titration of the enzyme (30). Other studies have also shown that this histidine, in particular, is instrumental in maintaining catalytic activity. Early experiments, using dyes and photooxidation, demonstrated the importance of such a residue in the enzymic activity of chymotrypsin (80). The elegant work of Mares-Guia and Shaw (1965) however, showed that specific alkylation

TABLE V

HISTIDINE 57 SEQUENCES

Protease	Sequence (57)
Trypsin-Bovine	Val-Val-Ser-Ala-Ala-His-Cys-Tyr-Lys-Ser-Gly-Ile-Gln
Trypsin-Porcine	Ala-Ala-His-Cys-Tyr-Lys
Trypsin-Turkey	Ala-Ala-His-Cys-Tyr-Lys
Trypsin-Shark (Spiny Dogfish)	Ala(Ala,His,Cys,Tyr,Arg,Ser,Gly)Ile-Gln-Val-Arg
Chymotrypsin A -Bovine	Val-Val-Ser-Ala-Ala-His-Cys-Gly-Val-Thr-Thr-Ser-Asp
Chymotrypsin A -Porcine	Ala-Ala-His-Cys-Gly-Val-Thr-Thr-Ser-Asp
Chymotrypsin B -Bovine	Val-Val-Ser-Ala-Ala-His-Cys-Gly-Val-Thr-Thr-Ser-Asp
Elastase-Porcine	Val-Met-Thr-Ala-Ala-His-Cys-Val-Asp-Arg-Glu-Leu-Thr-Phe-Arg
Thrombin-Bovine	Val-Leu-Thr-Ala-Ala-His-Cys-Leu-Leu-Tyr-Pro
SGT	Val-Leu-Thr-Ala-Ala-His-Cys-Val-Ser-Gly-Ser-Gly-Asn-Asn-Thr
SGPA	Val-Asn-Gly-Val-Ala-His-Ala-Leu-Thr-Ala-Gly-His-Cys-Thr-Ser
α-Lytic Protease-Myxobacter 495	Phe-Val-Thr-Ala-Gly-His-Cys-Gly-Thr-Val-Asn-Ala-Thr-Ala-Arg

of histidine-57 with TLCK was concomittant with loss in the enzyme activity of trypsin (31). Similarly, Beeley and Neurath (1968) demonstrated specific reaction of histidine-57 with bromoacetone (81). But even more important, they demonstrated that the reagent would react with histidine-57 in diisopropylphosphoryl-trypsin (DIP-trypsin) but that DFP would not react with the bromoacetone modified trypsin. This strongly indicated that bromoacetone, on reaction with histidine-57, destroyed the charge transfer system diagramed earlier, thus reducing the nucleophilicity of the active serine-195. One of the strongest corroborations of the implication of histidine-57 in the general reaction mechanism of all serine proteases, is the obvious conservation of this residue in the primary sequence. Table V is a compilation of sequences around histidine-57 in a selection of serine proteases (22).

The histidine-57 imidazole is important as an intermediary, through whose conjugated bonds electrons can be relayed between the charged aspartic acid-102 carboxyl group and the potentially nucleophilic serine-195 hydroxyl oxygen. Since the active serine-195 and histidine-57 residues have been discussed above, a brief discussion of aspartic acid-102 is now in order. The importance of aspartic acid-102 was deduced from X-ray diffraction measurements by Blow and co-workers (1969) (82). These workers observed that when histidine-57 is charged (below

TABLE VI

ASPARTIC ACID 102 SEQUENCES

Protease	Sequence (102)
Trypsin-Bovine	Asn-Asn-Asp-Ile-Met-Leu-Ile-Lys-Leu-Lys
Trypsin-Shark (Spiny Dogfish)	(Asn,Asp) Ile-Met-Leu-Ile-Lys-Leu-Ser
Chymotrypsin A-Bovine	Asn-Asn-Asp-Ile-Thr-Leu-Leu-Lys-Leu-Ser
Chymotrypsin B-Bovine	Arg-Asn-Asp-Ile-Thr-Leu-Leu-Lys-Leu-Ala
Elastase-Porcine	Gly-Tyr-Asp-Ile-Ala-Leu-Leu-Arg-Leu-Ala
Thrombin-Bovine	Asp-Arg-Asp-Ile-Ala-Leu-Leu-Lys-Leu-Lys
SGT	Gly-Lys-Asp-Trp-Ala-Leu-Ile-Lys-Leu-Ala
SGPA	Asn-Asp-Asp-Tyr-Gly-Ile-Ile-Arg-His-Ser
α-Lytic Protease-Myxobacter 495	Gly-Asn-Asp-Arg-Ala-Trp-Val-Ser-Leu-Thr

pH 6.0) the aspartic acid-102 and histidine-57 are close enough to be linked in an ion pair. Conversely, at those pH values greater than 6.0, where serine proteases are active, aspartic acid-102, histidine-57, and serine-195 are close enough to be involved in a hydrogen bond network extending from serine-195 to aspartic acid-102 via histidine-57 (80). Chemical modification of aspartic acid residues has been attempted with diphenyldiazomethane and glycineamide in the presence of a water soluble carbodiimide (80). The results indicated the importance of two carboxyl groups in chymotrypsin, one of which is aspartic acid-194, discussed previously and known from crystallographic work to be important for maintaining the enzyme in active conformation. Only indirect evidence for the role of aspartic acid-102 in the catalysis of the enzyme was obtained. Aside from the crystallographic evidence, one of the strongest arguments in favour of the importance of this residue in the general catalytic mechanism of serine proteases, is its conservation in the primary sequence of all known Asp-Ser-Gly proteases. This is shown in Table VI (22). The obvious high degree of similarity between sequences of enzymes from markedly different sources is of major importance in both defining the degree to which enzymes of a given group can be said to be chemically homologous and in defining the relative importance of markedly conservative residues in the general mechanism of action of the class of enzymes.

The concepts of chemical and conformational homology serve to define the structural relationships among the members of a group of proteins. As applied in this dissertation, chemical homology refers to the similarity in amino acid sequence of proteins which is greater than that which would be expected by chance alone. In many cases, sequence data have been exploited to reveal similar three-dimensional features and thus conformational homology, through the use of X-ray diffraction of protein crystals. Frequent reference will therefore be made to the work of Stroud et al (1971) and the description of the molecular structure of DFP inhibited bovine trypsin (35).

With specific reference to primary sequence data however, there are many regions of sequence within the serine proteases which display large amounts of homology and even identity. Such areas are the NH_2-terminal sequences, the disulphide bridge sequences, sequences close to the COOH-terminus and probably most important, the regions previously discussed involving the active site residues histidine-57, aspartic acid-102, and serine-195. Tables II, V, and VI depict selected sequences of peptides from proteases related to trypsin in the proximity of the active site residues (22). As mentioned previously, some of the SGT sequences have been extended beyond that published in the cited reference. There is little doubt that the sequences shown in the previous Tables indicate

a startling degree of homology within a taxonomic span represented by the most primitive proteases, characterized by both an actinomycete and a myxobacter, through to the most advanced mammalian proteases.

In the complete enzymes, the degree of homology and identity can be quantitated. For example, when the primary sequence of chymotrypsinogen A and trypsinogen are aligned, there is a coincidence of amino acids which represents 40% of the total sequence. Allowing for conservative replacements, the homologous areas can be extended to include 51% of the primary sequence. The degree of homology in chymotrypsinogens A and B is even more impressive. Approximately 80% of the residues in these molecules are identical, inferring that these two enzymes deviated from their common ancestor a shorter time ago than did chymotrypsinogen and trypsinogen. Data for these calculations was derived from the work of Smillie et al (1968) (83).

According to Perutz et al (1965), implicit in any discussion involving sequence homology is the tenet that extensive sequence similarity reflects a comparable similarity in tertiary structure (84). What this implies is that the information necessary for production of a competent tertiary structure is inherent in the primary sequence, and is not, for example, associated with some other outside parameter imposed at the time of protein

synthesis. The total synthesis of ribonuclease by Gutte and Merrifield in 1968, provides strong evidence for this theory (85). In agreement with this concept, Sigler et al (1968) reasoned that since the primary sequence homology between chymotrypsin and trypsin was so great, a model of the later enzyme could be constructed using the known tertiary structure of chymotrypsin as a backbone (57). Although there are two extra disulphide bridges in trypsin, when the model was built substituting trypsin residues for α-chymotrypsin residues, the two disulphide bridge bonds could be introduced with scarcely any movement of the main-chain atoms. According to the authors, this indicated that the two enzymes had nearly identical tertiary structures. This conclusion has been corroborated by the recent X-ray diffraction data of DIP-trypsin (35).

We have recently used the same principle as explained above, to construct a crude ribbon diagram of SGT, using α-chymotrypsin as a template molecule (86). This proved necessary as the X-ray diffraction coordinates for the tertiary structure of bovine trypsin were unpublished and remain so to this time. However, as will be shown later, the primary sequence homology between SGT and chymotrypsin is more than adequate to allow for this theoretical model and permits a comparison of a hypothetical tertiary structure of SGT with α-chymotrypsin and to a more limited extent, with the tertiary structure of trypsin.

Some of the more important primary sequence differences observed by earlier investigators have been considered in the context of the tertiary structures of trypsin and α-chymotrypsin by Stroud and colleagues (35). Many of the dissimilarities bear on the specificity differences of these two enzymes and are thus related to the specific binding site located very near the active site. Several of these variant positions have been discussed previously as they were relevant to the explanation of the significance of aspartic acid-189 in the specificity pocket. However, a more thorough discussion of the remaining differences is pertinent to the comparison to be made later between the ribbon diagrams of SGT and α-chymotrypsin.

As stated earlier, trypsin and its congeners demonstrate a high specificity for amino acid residues with charged basic side chains, while chymotrypsin is less specific, showing a preference toward residues with aromatic side chains and to a lesser extent, aliphatic side chains. In both cases the specific residue is on the NH_2-terminal side of the bond which is cleaved in catalysis. The unique trypsin specificity has long suggested an anionic site for the binding of substrates, and kinetic experiments have indicated that this is correct (60). As stated earlier, the anionic residue is known to be aspartic acid-189, a residue found in all 'trypsin-like' enzymes. This conclusion has been

corroborated by Stroud et al (1971) during X-ray diffraction studies using lysyl and arginyl amides (35). Stroud et al (1971) also observed that, in contrast to chymotrypsin which makes three hydrogen bonds to peptide substrates (87), trypsin appears to have the possibility of making four hydrogen bonds between peptide substrates as a result of a difference in the architecture of the binding pocket. These bonds are established on the NH_2-terminal side of the bond to be cleaved between the carbonyl and amino groups of the substrate peptide chain and those of serine-214, glycine-216, and glycine-219 of the trypsin main chain in an antiparallel β-pleated sheet arrangement. This phenomenon appears to be the result of a deletion in trypsin at residue 218.

An insertion in the trypsin sequence at position 221A provides a glutamine in the rear of the specific binding pocket which appears to form a hydrogen bond to aspartic acid-189. Stroud and colleagues presume that the function of this bond is to help freeze the conformation of this residue, decreasing the possibility of delocalizing the negative charge (35). Chymotrypsins possess a considerably different sequence in this region of the molecule between residues 221 and 224 (83).

A further region of difference between trypsin and chymotrypsin involves the insertions at 184A and 188A, which not only give rise to a larger specificity pocket

but, in the case of tyrosine-184A, helps to insulate the internal negative charge at the back of the cavity in the molecule (35).

These and other structural differences will be considered again when the comparison between the theoretical SGT tertiary structure and the tertiary structure of α-chymotrypsin is made in Chapter IV. Fundamental differences in sequence between SGT and other serine proteases will also be discussed.

Whereas the previously examined differences in primary and tertiary structures elucidate relationships between the structure and function of serine proteases, much information regarding the evolution of these enzymes can be obtained by consideration of the similarity or homology of their primary structures. Indeed, some of the most homologous regions of the serine proteases, observed about the active site residues, have been indicated already and their phenomenal similarity noted. Investigators have continually strived to recognize similarities and differences in living things which would allow for an acceptable theory accounting for the origin of the diversity of living forms. Indeed, using primarily morphological means, workers throughout the last hundred years have compiled overwhelming evidence to support the unitary theory of the origin of life, that is, that all living forms on earth, as we know them, are the

result of a single occurence.

Concurrent with this tremendous growth in morphological evidence was a growing biochemical knowledge allowing researchers a preliminary understanding of how function, structure, and information transfer were related. Emerging out of this, was a new field of study termed biochemical evolution, which has supplemented the morphological evidence and even, according to some workers, obtained direct definitive proof of the unitary theory (2, 88,89,90,91). Without question, several pieces of new information were instrumental in launching many laboratories into the new area of biochemical taxonomy. Perhaps the first was elucidation of the genetic code linking triplets of nucleotides in DNA and messenger RNA to individual amino acids (92) — a hypothesis put forth earlier by Crick (1958) (93). But researchers also required that there be a concise demonstration of the collinearity of base sequences and amino acid sequences in order to use the more readily obtainable amino acid sequence information for their work. This was established by the use of synthetic polypeptide sequences in the E. coli ribosomal system (94) and by comparing the genetic map of the E. coli tryptophan synthetase A locus, with the known primary amino acid sequence of the product protein (95). Since sequence analysis techniques for proteins had been well established long before this (96), work began immediately on the sequence of proteins which might,

as a result of their functional similarities, be structurally similar. But there were limitations to the protein sequence approach. One of the most serious was that, unlike the paleontologist, the biochemist had to work with contemporary structures. (Florkin (1964) however, has described a protein from Nautilus fossils which is presumed to be 6.5×10^6 years old, and has a similar amino acid analysis to the modern day equivalent protein (97)...). In face of this drawback, alternative approaches for comparative study have been taken in the absence of fossils. There have been those who have studied functionally related molecules within a single species and those who have chosen to study functionally related proteins up the phylogenetic tree. Both approaches use one or all of the following methods of comparative study:

1) Immunological techniques have been used. In most cases, once a purified antiserum is made, the test sample need not be purified to a high degree. This facilitates rapid screening but is relatively difficult to quantitate (91,98).
2) Enzymes can be subjected to kinetic analysis, specificity and stability studies, as well as physical studies, in order to provide evidence of taxonomic relatedness.
3) Sequence studies can be used in a rapid screening process by limited strategic labelling techniques, such as at

active sites, or used as a means towards constructing the total primary sequence of the molecule.

4) X-ray crystallography, generally the last step after primary sequence determination, leads to the most sophisticated comparative procedure, that of comparison of the tertiary structures of related molecules.

In general, the latter three methods appear to be used more than the first, but none can be discounted entirely.

Biochemists using the previously mentioned methods have visualized the evolution of proteins as having evolved from a primitive polypeptide which underwent random mutation and natural selection as the host "organism" evolved (99). At this point in prebiological time the primordial organism was emerging and with it, the greatest problem faced by evolutionists, that of the linkage of protein synthesis, energy metabolism and information transfer. This stage of association is lost due to a lack of information regarding transitional forms of organisms and must be left aside at this juncture. Many visionary reviews are available on prebiological evolution (100,101,102).

Before proceeding further it seems appropriate to quickly review the typical mutations observed in nucleic acids which inevitably result in the varieties of homologous proteins to be discussed. All mutations can be described as either point mutations or base sequence mutations,

where the latter refers to a mutation involving more than a single base. Point mutations can be subdivided into substitutions, deletions and insertions. The latter two mutations are the most deleterious as they result in a reading frame shift of one nucleotide. With respect to the more extensive base sequence mutations, these can be similarily subdivided. Sections of genome can be duplicated or deleted. The latter is believed to occur as one of two alternatives arising out of an inversion loop, where the sequence can simply invert or, as in the case of the deletion, be removed altogether. An excellent discussion of these 'chemo-genetic' manipulations was published in a review by Dixon (1966) (99).

All of the previously mentioned mutations are well documented in the literature, with perhaps the single point mutation in a normal form of human haemoglobin, producing the disease sickle cell anaemia, being one of the most obvious (103). At the other extreme, researchers believe that the various human non-allelic forms of the haemoglobins α^{A}, β^{A}, γ^{F} and δ^{A2} as well as myoglobin may be the result of whole gene duplications (104). Yet another example involves non-homologous pairing and cross-over which can result in partial gene duplication between non-allelic cistrons. This was elegantly demonstrated in the tryptophan synthetase system of Neurosporra crassa and Saccharomyces cerevisiae where tryptophan synthetase A

and B activities were observed in a protein which could not be broken into subunits. This is in contrast with the E. coli system where these activities are found in two separate proteins and are expressed therefore, by separate genes (105).

Since great amounts of sequence data are now being compiled certain immutable facts regarding these mutational events have become obvious. Comparison of the primary sequences of functionally similar proteases for example, has allowed for alignment of identical residues around key areas of the molecules. Those proteins having large areas of identical primary sequence were termed homologous proteins, as stated earlier. On examination of the non-identical residues in a group of homologous proteins by alignment of the identical regions, it was found that the non-identical residues could be grouped into two classes. Those that were structurally similar were said to be conservative substitutions while those that were structurally very different were said to be radical substitutions (88,106). For example, a valine substitution for an isoleucine would be considered conservative as compared with a valine for an arginine substitution. According to Epstein (1964) (106) using Kendrew's (1962) (107) myoglobin model, there is a central or core region of hydrophobic residues such as phenylalanine, leucine, isoleucine, valine, methionine, proline, and cysteine forming an area with a high dielectric

constant, while alternatively, according to Epstein, there are hydrophilic or polar residues such as lysine, arginine, glutamic acid, aspartic acid, glutamine, asparagine, serine, threonine, histidine, tryptophan, and tyrosine which are located on the outside adjacent to the highly polar aqueous phase. Residues such as tryptophan and tyrosine however, are known to be somewhat variable, depending upon the protein in question. Indeed, these designations are best regarded as generalizations, as on consideration of just the α-chymotrypsin molecule alone, tryptophan residues can be completely buried or entirely on the surface. Similarly, many of the very hydrophobic phenylalanine residues are observed to be completely exposed on the surface of α-chymotrypsin. In spite of these irregularities, if a non-polar residue such as phenylalanine, buried deep in the molecule, is substituted by an arginine as a result of a single point mutation, the consequence would almost surely be disruption of the tertiary structure of the molecule and loss of activity. Similarly, a very hydrophobic residue can be substituted for a polar residue on the surface of the protein only at a considerable expense energetically (88). Epstein's predictions regarding permissible or conservative replacements being far greater in number than the alternative non-permissible or radical substitutions, appear to be born out when one observes the replacements occuring in

sequenced proteins with known tertiary structures (88). It is interesting to note that even the genetic code seems predisposed towards conservative replacements as most non-polar residues for example, have the codon UU(Y), CU(Y), AU(Y), or GU(Y) where Y is any of the four possible nucleotides (99). Thus, all these hydrophobic residues are interconvertible at position one and have a bias of 2:1 toward conservative replacements among the non-polar residues. Conservative replacements and other mutational alterations in the proteases will be indicated later when their primary structures are compared in detail.

In summary, it should be stated that the bulk of the homology evidence from the primary sequence analyses of Asp-Ser-Gly proteases strongly indicates that this class of enzymes has evolved from a single ancestoral form through a divergent process analogous to the classical Darwinian descent with modification.

It is vastly important that not only structural considerations, but also functional data, be compiled in order that as much information as possible be gathered regarding the proteins being compared. From such data conclusions can be formulated, not only regarding the evolutionary homology of proteins, but also concerning those functions basic to biology, that the selective forces are acting upon to direct the design of the molecule.

The intent of this research was to investigate

the structural and physico-chemical properties of the serine protease SGT. Specifically, work was undertaken to elucidate the primary sequence structure of SGT along with those physical and kinetic parameters which bear directly on the problem of the functional and structural homology of the serine proteases. It was also hoped that the physico-chemical studies would help broaden the knowledge regarding the mechanism of action of the trypsins. In general, these goals have been attained, as the complete sequence of SGT has been elucidated, with the exception of one region containing a single residue overlap. Furthermore, the physico-chemical studies indicate that SGT is indeed enzymatically very similar to bovine trypsin, with the possible exception of the role of the NH_2-terminal α-amino group. It is felt that results arising from this research strongly indicate that the 'trypsin-like' enzyme, obtained from Streptomyces griseus pronase, is indeed homologous in structure and enzymic properties with all known trypsins and to a lesser extent with all serine proteases considered to be phylogenetically related to bovine trypsin.

CHAPTER II

GENERAL MATERIALS AND METHODS

1. MATERIALS

A. General Chemicals and Solvents

Bio-Rex-70, large pore size (200-400 mesh) and AG1-X2 (Dowex 1-X2) (200-400 mesh) were obtained from Bio-Rad laboratories, Richmond, California whereas Technicon Chromobeads Type P (peptides) equivalent to Dowex 50-X4 was obtained from Technicon Chemical Company, Chauncey, New York. CM-Sephadex, SE-Sephadex C-50 Lot #9757 (discontinued), SP-Sephadex C-25, Sephadex G-25 Fine, Sephadex G-50 Fine, Sephadex G-75 Fine and Sephadex G-75 Superfine were obtained from Pharmacia Fine Chemicals, Uppsala, Sweden. Ethylenimine was purchased from Dow Chemical Company, Freeport, Texas, redistilled before use and stored over sodium hydroxide pellets. Similarly, N-ethylmorpholine (practical grade) was purchased from Eastman Organic Chemicals, Rochester, New York and was redistilled prior to use (b.p. 134.0° C). BAEE (N-benzoyl-L-arginine ethyl ester), ATEE (N-acetyl-L-tyrosine ethyl ester), PNPA (p-Nitrophenyl acetate), oxidized insulin A and B chains and cacodylic acid were all purchased from Schwarz-Mann Chemical Company, Orangeburg, New York. Ultra-pure grade guanidinium hydrochloride and urea were also purchased from Schwarz-Mann. DFP (diisopropyl-fluorophosphate) was purchased from K & K Laboratories

of Hollywood, California, while cyanogen bromide was obtained from Pierce Chemical Company, Rockford, Illinois. Deuterium oxide was donated by Dr. M. Barton and was obtained from International Chemical and Nuclear Corporation of Irvine, California. TPCK (L-(1-tosylamido-2-phenyl) ethylchloromethyl ketone), TLCK (L-(1-tosylamido-2-lysyl) ethylchloromethyl ketone) and PAB (p-aminobenzamidine) were purchased from Sigma Chemical Company of St. Louis, Missouri. Dansyl amino acids were obtained from Calbiochem, Los Angeles, California or were prepared by the method of Gray (108). DNS-Cl (1-dimethylamino-naphthalene-5-sulphonyl chloride) B grade was also obtained from Calbiochem. Polyamide plates were purchased from Cheng Chin Trading Company, Taipei, Taiwan. Maleic anhydride (zone refined 99.9+%) was purchased from Aldrich Chemical Company, Milwaukee, Wisconsin. Durham tubes for dansylation were purchased from A. Gallenkamp and Company, London, EC2 England.

B. Enzymes

α-Chymotrypsin (three times crystallized), trypsin-TPCK, and pepsin (2600 U/mg) were purchased from Worthington Biochemical Company, Freehold, New Jersey, while crystalline thermolysin (B grade) and pronase (B grade) were purchased from Calbiochem, Los Angeles, California. Myxobacter 495 α-lytic protease was prepared in the laboratory of Dr.

D. R. Whitaker, University of Ottawa, Ottawa, Ontario.

2. METHODS

A. Column Chromatography Procedures for Preparation of SGT from Pronase

(1) Ion Exchange Gel Chromatography

Initial preparations of SGT followed precisely the procedure of Jurasek, Fackre and Smillie (1969) (109) and were generally found to be both laborious and wasteful of material. Several column procedures were required (CM-cellulose and Bio-Rex-70) as well as dialysis and lyophilization procedures after running each column. Since the mixture of enzymes and the purified SGT itself were known to autolyze under the conditions of this procedure, a new procedure was devised involving a one-step purification using volatile solvents. Sulphoethyl Sephadex (SE-Sephadex) C-50 was equilibrated with 0.02 M pyridine acetate buffer, pH 5.0 and packed in a 5 x 100 cm column which was equilibrated overnight with 0.02 M buffer and then loaded with between 10 and 20 g of pronase. It was found necessary to dialyze the crude pronase overnight at 5° against starting buffer and to centrifuge the product for 20 mins at 4000 x g before application to the column. The column was then developed with a linear gradient of pyridine acetate buffer from 0.02 M to 0.5 M at 5° with a gravity flow rate of 100 ml/hr. (molarities

are with respect to pyridine). The reservoir and mixing cylinder each contained 4.3ℓ of buffer. The volume of fractions collected was 25 ml which were read at 280 nm on a Gilford model 222A spectrophotometer while at 5°. The latter temperature is important as pyridine buffers have a temperature dependent extinction coefficient at this wavelength. Samples were pooled and lyophilized directly. An adaptation of this technique using CM-Sephadex and a linear gradient from 0.02 M to 0.75 M, with 6.5 liters in each cylinder, was later devised by Dr. L. Jurasek in our laboratory in order to even better resolve the SGT component at high loading levels (20). This procedure, together with a slightly modified SE-Sephadex chromatography utilizing a pyridine acetate gradient of 0.1 M to 0.5 M, with 4.3 liters of buffer in each cylinder, was routinely used to produce enzyme of high purity.

(2) Bio-Rex-70 Chromatography

The Bio-Rex-70 procedure of Jurasek et al (1969) was used (109). A column of Bio-Rex-70 resin, previously equilibrated with 0.1 N NaOH—cacodylic acid buffer, pH 6.10, was developed at 4° with the same buffer. The flow rate for a 5 x 100 cm column was 120 ml/hr. Fraction volumes of 20 ml were collected and the position of eluted protein determined spectrophotometrically.

B. Column Chromatography Procedures for Purification of Peptides

(1) Cyanogen Bromide Peptides Separated on Sephadex G-75 Superfine

Sephadex G-75 Superfine was swollen overnight in 0.65 N acetic acid, 8 M urea and packed in a Pharmacia column of either 2.5 x 100 cm or 5.0 x 100 cm and equilibrated overnight against the former solvent at room temperature. The flow rates were in the order of 8-12 ml/hr with a hydrostatic head of 90 cm. Sample sizes varied slightly but were approximately 25 mg in 5 ml and 600 mg in 30 ml for the 2.5 and the 5.0 cm columns respectively.

(2) Sephadex G-25 Fine and G-50 Fine Purifications of Tryptic Digests of Performic Acid Oxidized SGT

These gels were equilibrated and run in essentially the same manner as the G-75 Superfine columns, with the following exceptions: the buffer used was 0.5 M NH_4HCO_3, pH 7.8, the flow rates were 60 ml/hr and the fraction volume was 10 ml instead of 5 ml as used in the G-75 system. In one particular experiment the effluent from a G-50 column (5 x 100 cm) was fed directly into a G-25 column (2.5 x 100 cm) with all conditions remaining as above. These columns, unlike the G-75 Superfine columns were run at 5°.

TABLE VII

ELUTION SYSTEM FOR TECHNICON PEPTIDE ANALYZER

Chamber	0.1 M buffer,[a] pH 2.75 (ml)	0.2 M buffer, pH 3.0 (ml)	2.0 M buffer, pH 4.7 (ml)	2.0 M buffer, pH 6.5 (ml)	Water (ml)
1	90	—	—	—	—
2	90	—	—	—	—
3	90	—	—	—	—
4	—	90	—	—	—
5	—	90	—	—	—
6	—	55	15	—	20
7	—	5	45	—	38
8	—	—	87	—	—
9	—	—	65	22	—

[a] All buffers are pyridine-acetic acid and their molarities are with respect to pyridine.

(3) AG1-X2 (Dowex 1) Fractionation of Enzymatic Digests of Cyanogen Bromide Fragment Cn-1

The procedure of Landon (1964) for separation of peptides into basic, neutral and acidic fractions was used (110). A 2.5 x 15 cm column of AG1-X2 (200-400 mesh) equilibrated in 2% pyridine, 0.038% acetic acid buffer, pH 6.7 was equilibrated and peptides applied. Basic peptides were eluted with 60 ml of pH 6.7 buffer at 30 ml/hr and the fraction was lyophilized. The neutral fraction was eluted with 60 ml of 2% pyridine, 0.2% acetic acid buffer, pH 6, and the acidics containing some neutrals were eluted with 30 ml of 2 N acetic acid followed by 30 ml of 50% acetic acid. It should be noted that fractions I, II, and III did not contain only basic, neutral and acidic peptides respectively, but were crude fractions with some cross-contamination especially from the neutral fraction.

(4) Chromobead P Resin Fractionation of Basic, Neutral, and Acidic Peptides from Dowex 1 Column

The Technicon autoanalyzer system for peptide chromatography was used with a 0.6 x 100 cm column of Chromobead P resin. The column was jacketed and kept at 38° during operation. Elution was performed at 25 ml/hr using the buffer system employed by Welinder and Smillie (1972) as indicated in Table VII (111). A nine chambered

autograd was used to develop the gradient and three ml fractions were collected and located by automated ninhydrin assays before and after alkaline hydrolysis. All peptide containing fractions were pooled and lyophilized. This system utilized between 4 and 7% of the sample for analysis and gave good resolution on as much as 7.5 μmoles of protein digests. The basic system modified as just described is defined in greater detail in the Technicon Instruction Manual T-67-101 for peptide analysis.

C. High Voltage Paper Electrophoresis of Peptides

High voltage electrophoresis of peptides was routinely used as the final purification step in a protein digest isolation procedure. Electrophoresis at both pH 1.8 and pH 3.5 was performed in a Gilson High Voltage Electrophorator Model D equipped with a large fiberglass tank utilizing varsol as an inert coolant. This is basically the system described by Dreyer and Bynum (1967) (112). A third system was used at pH 6.5 and is essentially a vertical strip high voltage electrophoresis apparatus similar to that described by Ryle et al (1955) (113). Toluene, containing 8% pyridine (v/v), was used as a coolant for this system. Buffers used for all systems were as follows:

	Constituents	Ratios by volume
pH 6.5	Pyridine-acetic acid-water	(100:3:900)
pH 3.5	Pyridine-acetic acid-water	(1:10:189)
pH 1.8	Formic acid-acetic acid-water	(1:4:45)

Whatman 3 MM filter paper was routinely used for separation of peptides although Whatman No. 1 was sometimes used when only small quantities of peptides were available. For the preparative use of 3 MM, 50 nanomoles per cm of paper was considered the load limit beyond which yields and resolution suffered greatly. Electrophoresis was routinely carried out on full sheets of 3 MM for 45 min at 3 KV, which was equivalent to 60 volts per cm. Peptides were located on the dried sheets by utilizing cadmium-ninhydrin reagent (114) and when necessary the Pauly reagent for histidine (115), Ehrlich's reagent for tryptophan (116) and the α-nitroso-β-naphthol reagent for tyrosine (117). When a blocked NH_2-terminal was suspected, the starch-iodide technique for peptide bonds (118) was employed. The latter technique was modified by utilizing a weak solution of hypochlorous acid (commercially available as Javex) rather than the more reactive chlorine gas. All cadmium-ninhydrin treated papers were routinely heated at 60° for 15 min.

D. Amino Acid Analysis

(1) Peptide amino acid analysis

Amino acid analyses were routinely done on the Beckman Spinco Model 120 C automatic amino acid analyzer with expanded range card. Samples of peptides (30-50 nanomoles) were hydrolyzed in 100 µl of constant boiling HCl (5.7 N) for 16 to 24 hours at 110° in evacuated sealed tubes (10 mm x 75 mm). (All tubes were fired in a muffel furnace at 400° before use) The tubes were opened and the contents dried down in vacuo. In most cases 1/3 of the sample was placed on the short column, which resolves only basic residues, 1/3 was placed on the long column to resolve the neutrals and acidics, and 1/3 was retained for emergencies. Using the versatility of a programmable automated machine, all residues including S-β-aminoethyl-cysteine, cysteic acid, carboxymethylhistidine, homoserine and homoserine lactone were readily resolved from all the common amino acids with the machine operating constantly (119). The lower limit for quantitation on the expanded range card was approximately 0.5 nanomoles with a clean peptide and the upper limit was approximately 50 nanomoles, as resolution between closely running peaks became poor above this value. With the exception of serine, valine, tyrosine, isoleucine, aminoethylcysteine, homoserine, and homoserine lactone, which are either degraded, modified or poorly hydrolyzed, all amino acids were

obtained in acceptable molar ratios when peptides were analyzed using this procedure.

(2) Protein amino acid analysis

The amino acid composition of proteins and large protein fragments was determined on the automated Beckman Spinco Model 120 C amino acid analyzer with the 'protein' or normal range card. The procedure is identical with that used for peptides with the exception that approximately 0.10 μmoles of each amino acid in the protein was applied to obtain peaks of adequate size. Furthermore, samples were duplicated and hydrolyzed for 24, 36 and 72 hours to enable extrapolation back to maximum yields for degradable residues and to enable an equally valid assessment of slowly hydrolyzed residues such as valine and isoleucine.

E. NH_2-terminal and Sequence Determination of Peptides

The general procedure of Gray (1964) (108) as modified by Hartley (1970) (120) was used for NH_2-terminal determination and sequence of peptides. This 'Dansyl-Edman' technique allowed for minimum loss of peptide for NH_2-terminal determination and maximum speed of degradation. Two degradation steps could be accomplished in an 8 hour period and NH_2-terminal samples of approximately 1.0 nanomole were easily detected using the dansyl technique. The general procedure was as follows: approximately 0.1 μmole of peptide was taken up in 150 μl of water in a

ground glass stoppered tube and an appropriate volume removed for NH_2-terminal determination by dansylation.

(1) Dansylation and Identification of Dansyl Amino Acids

The sample removed for dansylation as described above was placed in a prefired Durham tube of approximately 4 x 25 mm. This sample was evacuated to dryness in a desiccator with a water aspirator, 5 µl of 0.2 M $NaHCO_3$ added and the sample again dried down by evacuation. Deionized distilled water (5 µl) and 2.5 mg/ml dansyl chloride in acetone (5 µl) were added to the sample which was incubated for 20 min at 45° while covered with Parafilm to prevent evaporation. After drying down (water aspirator), 25 µl of constant boiling HCl (5.7 N) was added, and the tube sealed by flame under vacuum. Samples were routinely hydrolyzed for 6 hours at 110° although valine and isoleucine containing peptides often required 16 hours. Tubes were cracked open by scoring and heating and the contents dried down under vacuum with an oil pump. Each sample was dissolved in 5 µl of acetone-acetic acid (3:2, v/v) and applied to double sided thin layer polyamide plates for thin layer chromatography. Approximately 1/3 of the sample was placed in a corner of the polyamide plate while another 1/3 was placed in the same corner of the opposite side along with standards. Plates were approximately 5 cm square and fitted easily into 250 ml beakers, where they

DIPEPTIDES SYSTEM 2

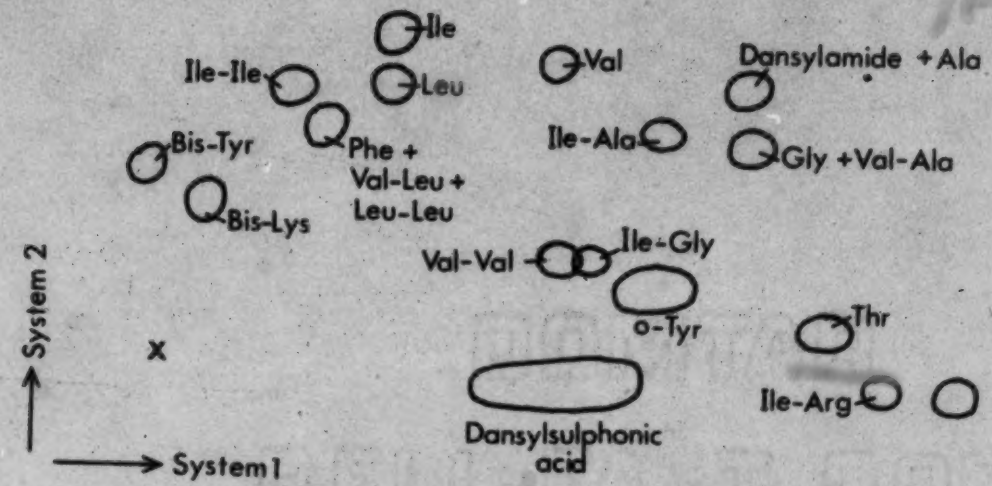


Figure 4

DIPEPTIDES SYSTEM 3

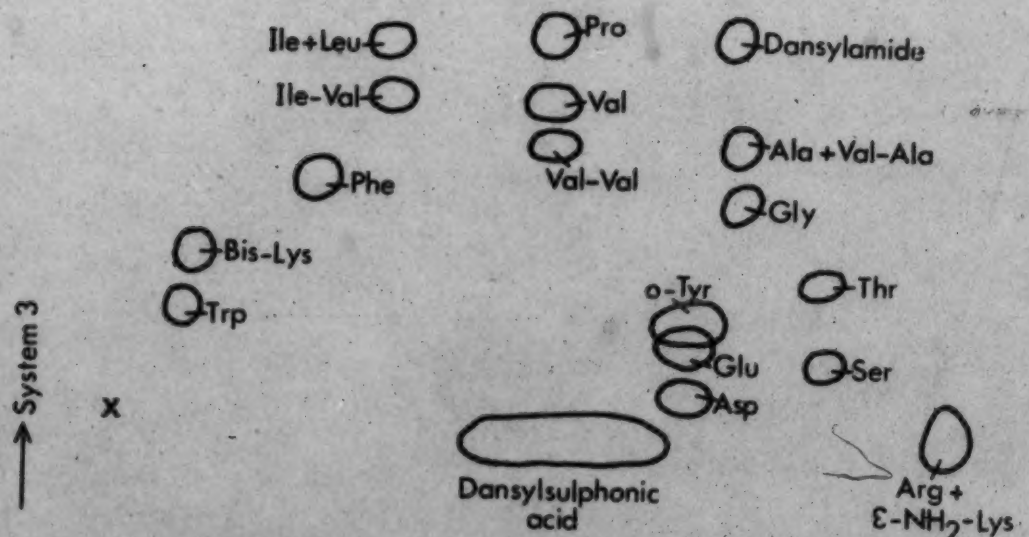


Figure 5

Fig. 4: See Facing Page.

Positions of some slowly hydrolyzing dansyl dipeptides, with standard dansyl amino acids, on a thin layer polyamide plate after chromatography in systems 1 and 2.

Fig. 5: See Facing Page.

Positions of some slowly hydrolyzing dansyl dipeptides, with standard dansyl amino acids, on a thin layer polyamide plate after chromatography in systems 1, 2 and 3. Systems 2 and 3 were run 90° to system 1.

SYSTEM 2

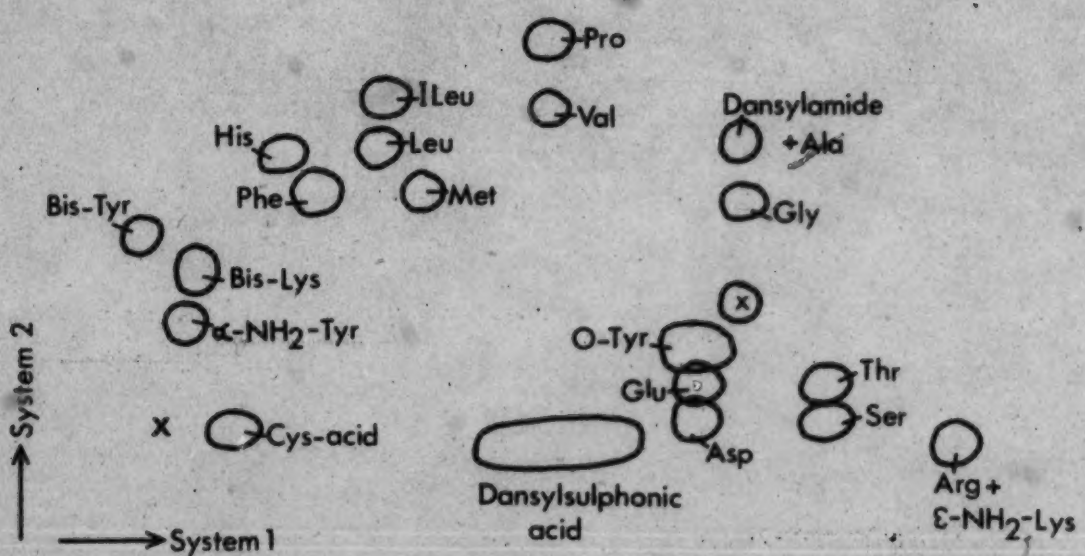


Figure 2

SYSTEM 3

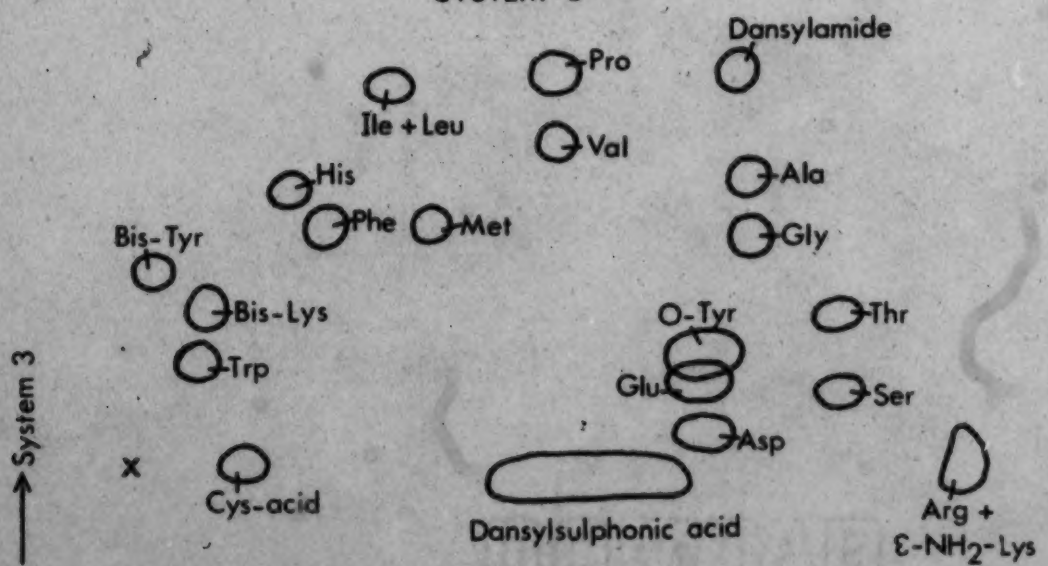


Figure 3

Fig. 2: See Facing Page.

Positions of dansyl amino acids on a thin layer polyamide plate after chromatography in systems 1 and 2.

Fig. 3: See Facing Page.

Positions of dansyl amino acids on a thin layer polyamide plate after chromatography in systems 1, 2 and 3. Systems 2 and 3 were run 90° to system 1.

were developed in two dimensions using the solvent systems employed by Hartley (1970) (120).

System 1: 1.5% formic acid (v/v)

System 2: Benzene-acetic acid (9:1, v/v)

System 3: Ethylacetate-methanol-acetic acid (20:1:1) by volume

A fourth system was often used to resolve dansyl-O-tyrosine from dansyl glutamic acid (121). It was n-butanol/n-hexane (1:1, v/v). Unequivocal identification of the NH_2-terminal amino acid was obtained after as many as sixteen degradations. Figures 2 and 3 describe the positions of dansyl amino acids after system 2 and system 3 respectively, while Figures 4 and 5 describe the positions of slowly hydrolyzed dipeptides after systems 2 and 3 respectively. Plates were washed in a solution of methanol—1 M ammonium hydroxide (1:1, v/v).

(2) Coupling with Phenylisothiocyanate (PITC)

The modified Edman degradation procedure described by Gray (1967), was used with certain alterations (122). To the remaining peptide sample 150 µl of 5% PITC in pyridine was added and the tube purged with N_2. (The 5% PITC was prepared fresh every two weeks and stored under N_2 in the freezer.) Tubes were placed in a 45° desiccator for 60 minutes for the coupling reaction and finally, in an evacuated 60° desiccator, with tops

removed, for 30 min in the presence of NaOH pellets and phosphorous pentoxide (P_2O_5). This procedure removes unreacted PITC, pyridine, water, and volatile by-products such as aniline and phenol, trapping them in an iso-propanol dry ice trap.

(3) Cleavage of Phenylthiocarbamylpeptide (PTC-peptide)

The above dried residue was dissolved in 200 µl of anhydrous trifluoroacetic acid (TFA) and flushed with dry N_2 to remove residual water vapour. Tubes were stoppered and incubated for 30 min at 45° in a heated dessicator and finally transferred to the 60° desiccator containing NaOH pellets and P_2O_5 where they were evacuated for 5 to 10 min.

(4) Removal of Diphenylthiourea

The contents of each Edman tube were dissolved in 200 µl of deionized distilled water and extracted three times with 1.5 ml aliquots of n-butyl acetate. Between each extraction the two phases were mixed thoroughly on a Vortex mixer and separated on a clinical centrifuge. The n-butyl acetate top layer was discarded and the remaining sample was taken to dryness at 60° over NaOH pellets and P_2O_5, while under an oil pump vacuum. Following the addition of 150 µl of deionized distilled water a sample was removed for dansylation and the peptide subjected to a further cycle of degradation.

F. Disc-gel Electrophoresis

The general procedures and materials used were those described and supplied by Canalco (Canal Industrial Corporation, Bethesda, Maryland) in their 1968 instruction manual for the Model 12 Canalco System. Since the isoelectric point of SGT was too high to use the pH 9.5 system, the pH 4.3 system with 7 1/2% acrylamide gel was used. This system stacks at pH 5.0 and runs at pH 4.3. Samples were run at 4 mA per tube and 12.5 volts per cm for 2 hrs with 5 ml of 0.01% methyl green (w/v) added to the top buffer compartment as a tracking dye. The bottom electrode was the cathode during the run but was reversed for destaining electrophoretically. The gels were stained for 2 hrs in a 0.5% Analine Blue Black solution which was 7% in acetic acid. Gels were also stored in 7% acetic acid.

G. Ultracentrifuge Studies

All molecular weight determinations were performed using a Spinco Model E analytical ultracentrifuge equipped with an electronic speed control. Only the low speed sedimentation equilibrium procedure was employed as described by Chervenka (1970) (123). These studies were monitored with a Rayleigh interference optical system according to Richards and Schachman (1959) (124). A multichannel sapphire windowed cell was used with concentrations of 1 mg per ml, 2 mg per ml and 3 mg per ml

SGT in 0.5 M KCl, 0.05 M sodium acetate, pH 5.0 buffer. Conditions of runs were as follows: Temp. 20°, overspeed time 3 hours, overspeed RPM 26,000, equilibrium speed 15,000, equilibrium time 18 hours. Other parameters used for calculations were the partial specific volume ($\bar{v}$)=0.717 cm^3/g as determined for human trypsin by Travis and Roberts (1969) (51), the solvent density (ρ_t) equal to 1.0244 and the refractive index increment $\frac{dn}{dc}$ = 0.185 (assumed). C_0, or the initial concentration of protein, was determined as accurately as possible by using a double sector synthetic boundary cell. Apparent molecular weights, MW_{app}, were calculated from the slope of ln y versus r^2 plots. (Where y=fringe displacement, r=distance from the axis of rotation, MW_{app}= apparent weight average molecular weight). All calculations were performed with the aid of an IBM 360 computer under the able direction of Dr. D. W. McCubbin with programs made available to us by Mr. W. T. Wolodko.

H. Optical Rotatory Dispersion and Circular Dichroism Measurements

Optical rotatory dispersion (ORD) and circular dichroism (CD) measurements were carried out in the laboratory of Dr. C. M. Kay by Mr. K. Oikawa whose skilled assistance is gratefully acknowledged.

a_o = complicated dispersion term which depends not only on the helical structure but also on environmental factors such as solvent, temperature and the nature of the residue side chains (128)

l = optical path length in dm

c = concentration in g/100 ml

m = average residue weight

n = refractive index of the solvent

Data computed in the usual manner with $-[m'] \lambda^2 - \lambda_o^2 / \lambda_o^2$ plotted versus $\lambda_o^2 / \lambda^2 - \lambda_o^2$ on the abscissa was used to evaluate protein solutions at each pH studied with respect to b_o which is the slope and a_o which is the intercept. The value λ_o was taken as 212 nm and the percentage α-helix was calculated from b_o, assuming that -640 degrees characterizes a fully coiled right handed α-helix (129).

The mean residue rotation values $[m']_{233}$, of -16,000° and -2,000° for completely α-helical and random coil polypeptides respectively, were used for calculation of percent α-helix at 233 nm.

For CD measurements the following equation was used to obtain the mean residue molecular ellipticity:

$$[\theta] = \frac{\theta m}{100\ lc}$$

where m = the mean residue molecular weight

c = the protein concentration in g/ml

l = the light path length of the cell in dm

A Cary Model 60 recording spectropolarimeter equipped with a Cary Model 6001 CD attachment and water-cooled lamp housing was used for these measurements in accordance with methodology described by Kay and Oikawa (1966) (125) and Oikawa *et al* (1968) (126). Samples of SGT (0.7 mg/ml) were subjected to visible (550-300 nm) and ultraviolet (250-200 nm) ORD scans at pH 5.0 through to 10.5 at 10°. The solvent system used was 0.05 M KCl titrated to the desired pH with 0.1 N HCl or 0.1 N NaOH. All solutions with pH values greater than 8.0 contained 20 mM glycine. CD scans (250-190 nm) used the same solvent system but protein concentrations were only 0.2 mg/ml. In ORD experiments the α-helix content was calculated for each pH value studied using the Moffit equation (127). This equation is as follows:

$$[m'] = \frac{a_o\lambda_o^2}{\lambda^2-\lambda_o^2} + \frac{b_o\lambda_o^4}{(\lambda^2-\lambda_o^2)^2}$$

where

$[m'] = \frac{3}{n^2+2} \quad \frac{m}{100} \quad \alpha$ = reduced mean residue rotation

$[\alpha] = \frac{100\alpha}{lc}$ = specific rotation

α = observed rotation

λ = wavelength at which rotation is measured

b_o = a constant insensitive to environmental changes and essentially an intrinsic property of the helical skeleton (127)

$[\theta]$ = the mean residue molecular ellipticity in degrees $cm^2/dmole$

θ = observed ellipticity

Reference values used for determination of percent α-helix came from the work of Greenfield and Fasman (1969) and are inherent in the expression below (130).

$$\% \alpha\text{-helix} = \frac{[\theta]_{208nm} - 4,000}{33,000 - 4,000}$$

I. Enzyme Assays

(1) Spectrophotometric Determination of N-benzoyl-L-arginine ethyl ester (BAEE) Hydrolysis by SGT

The hydrolysis of 1 mM BAEE, in a total reaction volume of 3.1 ml containing 10 mM $CaCl_2$, 50 mM Tris/HCl at pH 7.5, was performed at 254 nm with a temperature of 30° on a Zeiss single beam recording spectrophotometer (PMQ II indicating instrument and M4 Q III monochromator with a Sargent Model SRI Recorder)(131). The spectrophotometer was zeroed against the substrate solution and the final optical density, after complete reaction with SGT, was approximately 0.920 in a 1 cm cell (dilution due to addition of 100 µl of enzyme solution was taken into account). One unit of enzyme activity was that activity which would hydrolyze 1 µmole of substrate per min. Specific activity (SA) was defined as units of enzyme

activity per mg of SGT. The relationship between SGT concentration and $A_{280}^{1\ cm}$ was 1.72 optical density units = 1 mg/ml.

(2) The pH-stat Determination of N-acetyl-L-tyrosine (ATEE) and BAEE hydrolysis by Pronase Enzymes

Esterase activity was measured in a pH-stat with a Radiometer TTTla titrator and Radiometer Titrigraph Type SBR2c recorder with a temperature controlled water jacket and nitrogen-purged reaction vessel essentially according to the procedure of Walsh and Wilcox (1970) (132). Assay conditions were the following: 10 mM BAEE or ATEE, 20 mM $CaCl_2$, 0.1 M KCl, 10 mM Tris/HCl, pH 8.0 with a total volume of 3 ml maintained at 25°. The enzyme volume was usually 100 µl. Enzyme units and specific activity were expressed as stated earlier. When D_2O (deuterium oxide) was used the pH^2 was assumed equal to 'pH' + 0.40, where 'pH' was the reading on the pH-meter when D_2O was the solvent (133). All titrations were done with standardized sodium hydroxide (0.2 M) at pH 8.0 unless otherwise specified in results. Sodium deutroxide (NaOD) was made from D_2O and NaOH for the heavy water experiments.

(3) The pH-stat Determination of Oxidized Insulin A and B Chain Hydrolysis by Trace Contaminants in SGT

Conditions were as above except that 10 mM sodium hydroxide was used and the pH-stat was set to titrate at pH 9.0. Oxidized chains were assayed separately under the following conditions: 2.5 µmoles of substrate in 3 ml of standard pH-stat solvent (see above) and 25 µl of 10 mg/ml SGT (SA = 128 Units/mg).

(4) Spectrophotometric Determination of PNPA (p-nitrophenyl acetate) Hydrolysis by Pronase Enzymes

The procedure followed was essentially that of Wählby (134). The reaction was followed on a Beckman DBG double beam recording spectrophotometer at room temperature. Substrate was prepared by diluting 1 ml of a stock solution of absolute methanol, 35.4 mM in PNPA, with water to 100 ml. A 1 ml portion of this solution was then added to 1.9 ml of 50 mM sodium phosphate buffer, pH 7.0 in a 1 cm path length cuvette. An equal volume was added to 2.0 ml of buffer in the reference cuvette. The reaction was started immediately after thermal equilibrium was established by adjusting the machine to 400 nm and adding a 100 µl volume of enzyme solution. Results were expressed in nanomoles hydrolyzed/min/ml of enzyme solution added (20).

(5) Assay for Carboxypeptidase A Activity in SGT

The procedure used was a modification of the method of Narahashi and Yanagita (1967) (19) as detailed by Jurasek et al (1971) (20). The substrate was 20 mM carbobenzoxyglycyl-L-leucine (CbzGly-Leu) in 0.2 M sodium borate buffer pH 7.6. After addition of 0.1 ml of enzyme, reaction was allowed to proceed at 37° for 15 min and stopped by addition of 0.6 ml of ninhydrin reagent. After heating in a boiling water bath for 15 min, 2 ml of 50% ethanol was added and the fractions were read at 570 nm on a Gilford model 222A single beam spectrophotometer. The ninhydrin reagent (135) was made up during incubation by adding 1 g of ninhydrin and 0.15 g of hydrindantin to 37.5 ml of methyl cellosolve. This was allowed to dissolve and 12.5 ml of 4 N sodium acetate buffer pH 5.5 was added.

J. Digestive Procedures with α-Lytic Protease, Trypsin, α-Chymotrypsin and Pepsin

(1) Digestion of Cyanogen Bromide Fragment Cn-1 from SGT with α-lytic Protease

The substrate, Cn-1, is an S-aminoethylated cyanogen bromide fragment found to be very metastable with respect to solubility in most aqueous systems. Therefore, in order to solubilize this large peptide (molecular weight approximately 16,000) 12 ml of 50%

acetic acid was added to 300 mg (18.8 μmoles) of lyophilized peptide, and on dissolution, 60 ml of water was added. (Urea was found unnecessary when 50% acetic acid was used.) The temperature was maintained at 31°. The pH was adjusted to 8.2 with 29 ml of 6 N NH_4OH after addition of 3 ml of N-ethylmorpholine, and 3.6 mg of the protease was added immediately. The enzyme was added to the solution very rapidly after titration as the peptide was found to come out of solution at this pH, over a short period of time. Reaction was followed turbidimetrically at 660 nm until 40 min had transpired. Another 3.6 mg quantity of α-lytic protease was added at this time and the digestion was finally stopped after 128 minutes by freezing the sample. This final addition of enzyme brought the molar enzyme to substrate ratio to 1:50. An insignificant precipitate remained after centrifuging the sample prior to lyophilization.

(2) Tryptic Digestion of Cn-1 Cyanogen Bromide Fragment from SGT

Prior to digestion, 3.4 μmoles of Cn-1 fragment was dissolved in 3 ml of 50 mM acetic acid, 8 M urea and dialyzed at 8° for 48 hours against 20 liters of 50 mM acetic acid in a constant flow dialysis apparatus. After adjustment of the solution to pH 8.0 by the addition of 0.1 N NH_4OH, considerable precipitation occurred. Digestion was initiated by the addition of 1 ml of a 1 mg/ml solution

of TPCK-trypsin to the 4.5 ml volume at 25°. The finely dispersed suspension did not appear to clear during the 3 hr incubation time, although digestion did occur as NH_4OH (0.2 M) was added by the titrator throughout the reaction period. The reaction was terminated by addition of acetic acid to below pH 4.0. The precipitate was removed by sedimention in a clinical centrifuge and the pellet washed twice with small aliquots of 50 mM acetic acid. Both pellet and pooled supernatants were retained—the latter being applied directly onto paper for electrophoretic separation.

(3) Tryptic Digestion of Performic Acid Oxidized SGT

After performic acid oxidation (this chapter, section K7) 238 mg (10 μmoles) of SGT was dissolved in 100 ml of 50 mM N-ethylmorpholine acetate, pH 8.0—8 M urea and immediately dialyzed against 26 ℓ (13 ℓ x 2 changes) of 50 mM N-ethylmorpholine acetate, pH 8.0 for 24 hours at room temperature. To this stirred solution (149 ml) was added 2.13 mg of trypsin (TPCK) and the digestion allowed to proceed for 4 hours at room temperature. The reaction was terminated by freezing and lyophilization.

(4) Chymotryptic Digestion of Cn-1 Cyanogen Bromide Fragment from SGT and Maleylated Performic Acid Oxidized SGT

a. Digestion of Cn-1

Maleylated, S-aminoethylated Cn-1 (200 mg) was dialyzed free of maleic acid and urea, against 50 mM N-ethylmorpholine acetate, pH 7.5, at room temperature. This 500 ml solution was clear on initiation of the digestion but became slightly opalescent as the reaction proceeded. Conditions for digestion were as follows: 200 mg (12.5 μmoles) of Cn-1, 50 mM N-ethylmorpholine acetate, pH 7.5, 28°, with 10 mg of α-chymotrypsin (1:22 Molar Ratio) for 37 hours. The digestion mixture was continually stirred till the reaction was halted by addition of concentrated formic acid to pH 3.5. The sample was immediately placed in a 60° desiccator for demaleylation as described in this chapter, section K4.

b. Digestion of Maleylated Performic Acid Oxidized SGT

The digestion was started by adding 13.0 nanomoles of α-chymotrypsin to 2.7 μmoles of maleylated performic acid oxidized SGT dissolved in 5 ml of 1.0 mM N-ethyl-morpholine acetate, pH 8.0. (1:200 molar ratio enzyme to substrate). Reaction was allowed to proceed at room temperature with automatic addition of 0.2 M NH_4OH by the titrator to maintain a pH of 8.0. The material appeared clear when the reaction was terminated by lyophilization

after five hours.

(5) Peptic Digestion of TLCK-SGT

A typical digestion was done in 4.4 ml of 5% formic acid containing 23 mg (1 μmole) of inhibited SGT and 2.16 mg of pepsin. The sample was reacted in a 10 ml round bottom flask for easy lyophilization and incubation was carried out overnight at 37°. The digest was lyophilized directly and was found difficult to redissolve in glacial acetic acid. However, when applied to Whatman 3 MM over 25 cm and electrophoresed at pH 6.5 all of the material appeared to dissolve in this buffer.

K. Chemical Modification of Proteins and Peptides

(1) Cyanogen Bromide Cleavage of SGT

The method of Gross and Witkop (1962) was used with a molar excess of reagent to methionine residues of 200:1 respectively (136). The reaction consisted of 68.3 ml of 70% formic acid containing 397 mg (17.3 μmoles) of SGT and 0.444 g of cyanogen bromide. After 18 hours at room temperature a ten fold excess of water was added and the protein solution was lyophilized.

(2) Reduction and S-β-Aminoethylation of Cyanogen Bromide Fragmentation Peptides from SGT

The procedure of Raftery and Cole (1966) was used as described below (137). The cyanogen bromide cleaved product (800 mg) was dissolved in 120 ml of urea-Amediol-EDTA buffer (7.5 M urea, 0.1 M Amediol (2-amino-2-methyl-1,3-propanol), 0.3 mM EDTA), pH 9.5 and the protein fragments allowed to denature at room temperature for one hour, at which time 0.8 ml of mercaptoethanol was added. This solution was allowed to stand overnight in a N_2 atmosphere. Ethylenimine was added to the stirred reactant solution under a N_2 atmosphere while being automatically titrated at pH 9.0 with concentrated HCl. Addition of reagent was made in approximately 1 ml aliquots over 165 minutes till a total of 16 ml was added. At this time the rate of HCl uptake had markedly decreased. The product solution was unclear after reaction and further precipitation occurred on subsequent dialysis against 0.65 M acetic acid—8 M urea followed by 0.65 M acetic acid alone to eliminate urea. This stepwise dialysis was found necessary to obviate the sudden and much larger precipitation following immediate dialysis against acetic acid.

(3) Maleylation of Amino Groups in Performic Acid Oxidized SGT and S-Aminoethylated Cn-1 Cyanogen Bromide Fragment

The general procedure of Butler et al. (1969) was used (138). Reaction was carried out with the pH-stat in 96 ml of 0.2 M sodium borate buffer, pH 9.0, with 2.7 μmoles of performic acid oxidized SGT. Maleic anhydride was dissolved in 1,4-dioxane to make a 1 M solution. The pH was automatically adjusted to pH 9.0, after each addition of reagent, by 1.0 N NaOH. Additions of 1.6 ml were made every 5 min for 20 min. The performic acid oxidized enzyme was initially insoluble in this buffer system but was solubilized with maleylation.

Maleylation of 200 mg (12.5 μmoles) of the Cn-1 cyanogen bromide fragment was performed in 320 ml of 25 mM sodium borate buffer, pH 9.0, 50 mM mercaptoethanol (to reduce carbamylation) and 8 M urea. Titration at pH 9.0, was performed automatically with addition of 1 N NaOH, as 2 ml aliquots of 2,4-dioxane, 1 M in maleic anhydride, were added every 5 minutes for 50 minutes. The clear solution was finally dialyzed against two 2 ℓ volumes of 50 mM N-ethylmorpholine acetate, pH 7.5 for 18 hours at room temperature.

(4) Demaleylation of Amino Groups of Samples in Solution and on Paper

Demaleylation of maleylated polypeptides and proteins was done as suggested by Butler et al (1969) where the conditions are essentially pH 3.5 at 60° for 6 hours (138). In the case of peptides on paper, which were demaleylated to take advantage of the diagonal technique worked out by the Hartley laboratory (138), a 60° desiccator was utilized and the guide strips to be treated were wound about a glass spiral and inserted, along with a small volume of pH 3.5 pyridine acetate buffer, into the desiccator. A weak vacuum was applied with a water aspirator and the desiccator was incubated for 6 hours. The paper strips were then dried by evacuation in another desiccator in preparation for the second dimension of the diagonal procedure.

The procedure for solutions is identical with the exception that the solution was usually brought to pH 3.5 with formic acid and the flask closed with a teflon stopper.

(5) Acetylation of SGT

The procedure of Oppenheimer et al (1966) was followed with the exception that 3.0 mg/ml SGT was used instead of 10 mg/ml as used by the former authors (56). It was thought that a decreased concentration might mitigate any tendency for the acetylated enzyme to

aggregate. Unfortunately, this was not the case. The reaction was performed in a pH-stat titrating with 5 N NaOH while maintaining the temperature at 4° and the pH at 8.0. To 12 mg of enzyme in 4.0 ml of water was added five 10 µl additions of acetic anhydride over 40 min. The sample appeared opalescent after 30 µl was added and continued to precipitate thereafter. The insoluble suspension was then dialyzed overnight at 0° against 4 ℓ of water titrated to pH 4.0 with acetic acid. The precipitate and supernatant were then separated by centrifugation.

(6) Carbamylation of the NH_2-Terminus of SGT with Potassium Cyanate

In an attempt to specifically block the NH_2-terminal α-amino group in SGT, it was thought possible that the different pK_a's of the α-amino and ε-amino groups (approximately 8.0 and 10.5 respectively) could be taken advantage of, as it is known that α-amino groups will react with potassium cyanate (KNCO) approximately 100 times faster than ε-amino groups at neutral pH (139). As a consequence, the α-amino group might be selectively modified.

The procedure of Rimon and Perlmann (1968), for the carbamylation of the α-amino terminal leucine of pepsin, was used (140). The enzyme was brought into

solution in 5 ml of water in the reaction vessel of the pH-stat and 0.4 g of KNCO was added. The pH was maintained at 7.0 by automatic titration with 6 N HCl during rapid mechanical stirring and the temperature maintained at 0°. The soluble product, after 150 min of reaction, was dialyzed against 12 ℓ of 50 mM acetic acid, pH 3.5 at 0° for 20 hrs and finally lyophilized.

(7) Performic Acid Oxidation of SGT and Peptic Peptides of SGT on Paper

Performic acid oxidation of SGT was performed according to the method of Hirs (1967) (141) for cleavage of disulphide bonds. In general, a 10 fold excess of reagent over the quantity required to achieve oxidation of the sulphur containing residues was used. A typical reaction utilized 1 μmole of SGT dissolved in 0.32 ml of methanolic formic acid (1 ml methanol: 5 ml 99% formic acid) and 0.53 ml of performic acid reagent at -7° (salt ice bath). The reaction was terminated after 2.5 hours by dilution with 25 ml of water at 0° and immediately lyophilized. The performic acid reagent was prepared by incubating a mixture of 1 ml H_2O_2 with 19 ml of 99% formic acid at 25° for 2 hours.

Performic acid oxidation of peptic digests of SGT was performed while the peptic peptides were still on guide strips of electropherograms. This technique

was used in conjunction with the diagonal mapping procedure for detection of disulphide bond peptides. Guide strips were simply subjected to performic acid fumes from the above mentioned solution in a closed, weakly evacuated desiccator for 2 hours. After this time the strips were dried over NaOH pellets in vacuo for at least 1 hour.

(8) Active Site Directed Inhibition of SGT by TLCK

The reagent TLCK was used according to Shaw (1967) to inactivate SGT (142). This procedure was used exactly as described by Shaw for the preparation of inactivated enzyme to be used for chemical characterization. For studies of the rate of inactivation the procedure was modified in the following manner. The buffer used was 50 mM tris maleate, pH 7.0 containing 5.0 mM $CaCl_2$. An enzyme solution of 2.5 mg/ml and inhibitor solution of 0.25 mg/ml were made up and the reaction started by mixing both and incubating at 30° in a circulating water bath. Samples were taken with time and assayed immediately using the standard spectrophotometric BAEE assay previously discussed. Samples of 10 µl were taken and the reaction was virtually complete in 30 minutes. The inhibited enzyme was dialyzed against water overnight and lyophilized.

L. Diagonal Procedure for Identification of Disulphide Bond Peptides

The procedure used was that of Brown and Hartley (1963) and involved an initial pH 6.5 high voltage electrophoresis of peptic peptides of SGT prepared as described in this chapter, section J5 (143). The digest was streaked onto Whatman 3 MM paper and electrophoresis was performed for 45 min at 60 volts per cm. Side strips were cut from the electropherogram and stained with cadmium ninhydrin to detect the peptide positions. A further strip was removed for performic acid oxidation as described in this chapter, section K7. The oxidized guide strip was then sewn onto another sheet of Whatman 3 MM at right angles to the previous direction of peptide movement and electrophoresis was again performed at pH 6.5 for 45 min at 60 volts/cm. The cysteic acid peptides, produced by oxidation of the disulphide peptides present in the initial electropherogram, were found to lie off the diagonal produced by all other peptides whose mobility had not been altered by the oxidation.

M. Specificity Experiments with Oxidized Insulin A and B Chains

All experiments were conducted using SGT which had been purified on CM-Sephadex, followed by chromatography on SE-Sephadex and finally Bio-Rex-70. The specific

activity of the enzyme was 154 Units/mg SGT. Digests were done with enzyme to substrate molar ratios of 1:20 and 1:1000 in 1 ml of 50 mM N-ethylmorpholine acetate buffer, pH 7.5 for 1 hour at room temperature. Substrate concentrations were 0.125 μmoles/ml for all experiments. Enzyme inhibited with TLCK was prepared as discussed in this chapter, section K8, while enzyme inhibited with TPCK was done according to the procedure of Stevenson and Smillie (1968) (144). (To 0.1 μmoles of SGT dissolved in 3.1 ml of 0.1 M Tris buffer containing 50 mM $CaCl_2$ at pH 7.5 was added 100 μl of ethanol followed by 0.4 mg (1.0 μmole) of L-TPCK in 60 μl of ethanol to give a molar excess of TPCK to enzyme of 10:1. Reaction time was 2.5 hrs at room temp). The previous experiments were carried out to test the effects of inhibitors and various enzyme to substrate ratios, using small quantities of substrate. The procedure used to prepare peptides in larger amounts involved digesting 1 mg A chain (0.5 μmoles) and 3 mg B chain (1 μmole) at a 1:20 molar ratio and 1:1000 molar ratio. Digests were streaked on Whatman 3 MM and peptides were isolated electrophoretically as described in this chapter, section C. All isolated peptides were subjected to amino acid analysis as described in section D1 of this chapter.

N. Automated Sequence Determination of SGT

The Beckman 890 B Sequencer was used to verify the sequence of SGT for the first 55 residues. Procedures used were essentially those of Edman and Begg (1967) (145). The sample was brought into solution with 0.35 ml anhydrous trifluoroacetic acid and dried down in the spinning reaction cup. This material was then washed with 1-chlorobutane containing 0.005% 1,4-butanedithiol (v/v), dried in vacuo and subjected to the normal automated sequential Edman degradation using Quadrol buffer (N,N,N',N'-tetrakis-(2-hydroxypropyl)-ethylenediamine). The 2-anilino-5-thiazolinone derivatives were too unstable for identification purposes and were converted to PTH derivatives (3-phenyl-2-thiohydantoins) by incubating in 1 N HCl at 80° for 10 minutes. All samples with the exception of PTH aspartic acid, glutamic acid, lysine, asparagine, glutamine, isoleucine and leucine were injected directly into a Beckman GC 45 gas chromatograph fitted with 1.2 m x 2 mm, (inside diameter) glass columns packed with 10% SP-400. Approximately 10% of the sample was applied to the GLC which uses a helium carrier flow rate of 60 cc per min and a hydrogen flame detector system. The column was developed under the following conditions: 2 minutes isothermal at 165° followed by a 16 min rise phase through 110° to 275°. The remaining previously mentioned residues were treated with BSA (N,O-bis-(trimethylsilyl)-acetamide) for conversion to the silylated derivatives.

Arginine was identified qualitatively by the phenanthrene-quinone method of Yamada and Itano (1966) (146) and histidine was determined with the Pauly reagent (114).

O. Determination of the Molar Extinction Coefficient of SGT

Two samples of highly purified SGT, 3 mg each, were dissolved quantitatively in 3 ml of deionized distilled water at 4°. Two 1 ml samples were taken from each to be hydrolyzed at 22 and 96 hours for amino acid analysis. The remaining milliliter in each duplicate was used for optical density readings by diluting each sample to 0.4 mg/ml. The final optical densities were read in a Cary Model 15 recording spectrophotometer, Zeiss Model PMQ II-M4 QIII spectrophotometer, Gilford Model 240 spectrophotometer and Beckman DB-G recording spectrophotometer and the results averaged. The values for alanine and leucine obtained from the amino acid analyses were used in conjunction with the known amino acid composition of the protein to calculate the concentration of the solutions. The extinction coefficient could then be calculated directly.

P. Stability Studies with SGT in the Presence and Absence of Ca^{2+}

All studies were carried out at a protein concentration of 0.2 mg/ml in the following buffer systems:

pH 1.0, 2.0	HCl/KCl 10 mM	in KCL
pH 3.0, 3.5	sodium formate	10 mM in formate
pH 5.0	sodium acetate	10 mM in acetate
pH 6.0-8.0	tris maleate	10 mM in tris
pH 9.0-11.0	sodium glycinate	50 mM in glycine

The SA (specific activity) of the SGT was 143 U/mg SGT. When calcium chloride was added it was present at a concentration of 15 mM. All assays were done using 100 µl samples of incubated enzyme in the standard spectrophotometric assay detailed previously with the exception that the pH was 8.0 instead of 7.5. All pH values were read before and after incubation. Incubations were maintained at 0° in a refrigerated ice bath or at 30° in a circulating water bath. Since addition of 100 µl of sample to the standard assay solution constituted a thirty fold dilution with 50 mM tris/HCl, the pH of the assay mixture was essentially unchanged and all assays were therefore conducted at pH 8.0.

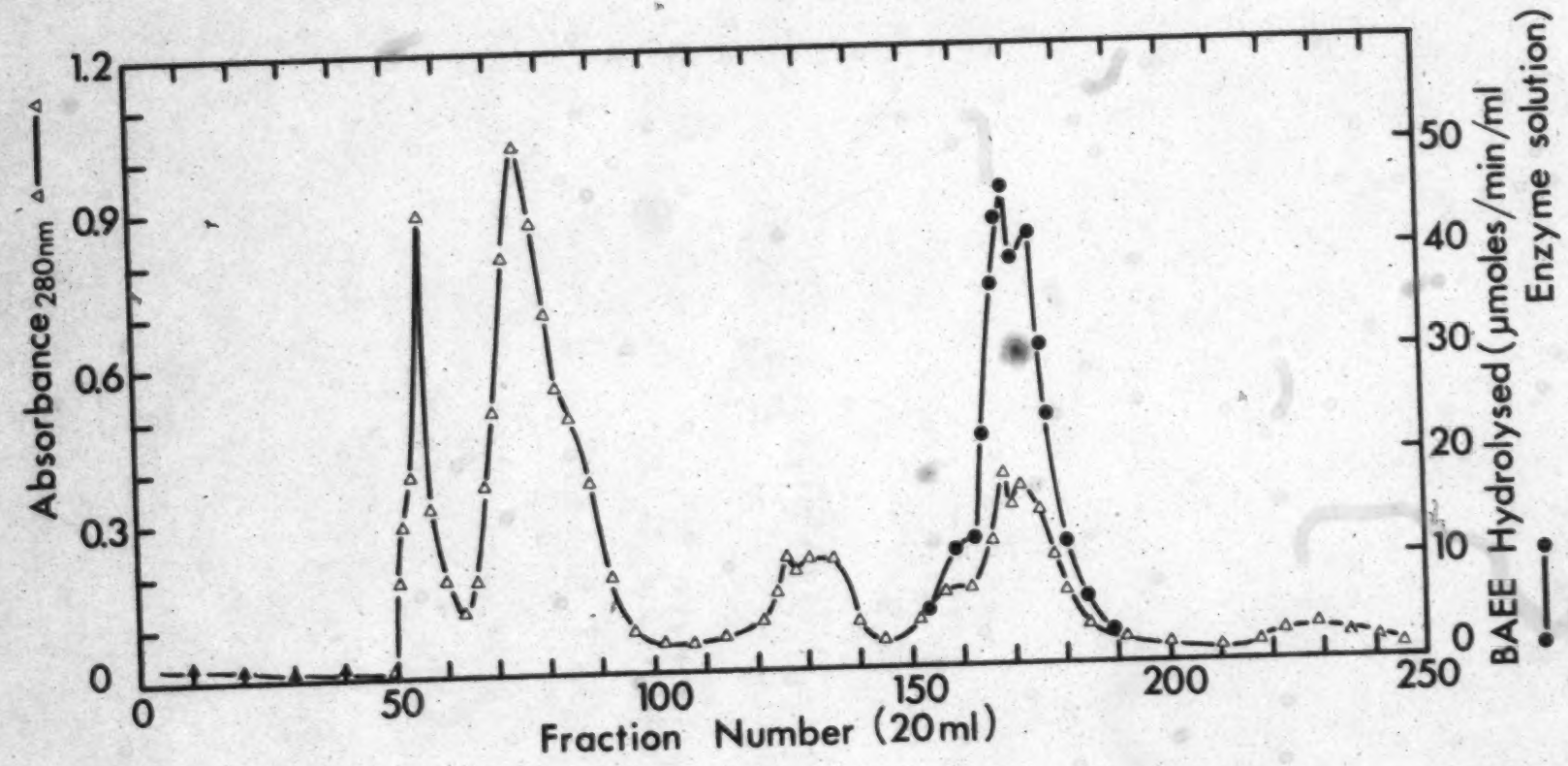


Fig. 6: Bio-Rex-70 chromatography of 275 mg of pronase previously fractionated on CM-cellulose. The sample was applied to a 4.5 x 65 cm column equilibrated with 0.1 N NaOH-cacodylic acid buffer, pH 6.10. Fractions were collected at 60 ml/hr.

CHAPTER III

RESULTS

1. PHYSICO-CHEMICAL RESULTS

A. Preparative Work

Initial attempts at preparation of purified SGT were made utilizing material which had been chromatographed on CM-cellulose (carboxymethyl-cellulose) according to the method of Wählby and Engstrom (1968) (16). However, this method produced a heterogenous product which necessitated the use of additional separation procedures to obtain enzyme of acceptable purity. Furthermore, fractions eluted from CM-cellulose contained calcium salts and required extensive dialysis before lyophilization. The impure fraction was applied to a Bio-Rex-70 column of dimensions 4.5 x 65 cm and eluted with a 0.1 N sodium hydroxide-cacodylic acid buffer system at pH 6.1 (109). A typical profile of such a preparative run on Bio-Rex-70 is shown in Figure 6. It should be noted that an assymetrical peak was observed which followed the activity profile and was assumed to be the result of autolysis products. Amino acid analyses performed on both peaks were indistinguishable within the limits of the technique. The yield of SGT using this two column preparative procedure, beginning with Pronase B grade, was 1%.

(1) SE-Sephadex C-50 Chromatography of Pronase

Since several other workers had reported CM-cellulose chromatography procedures which gave no increase in resolution, it was decided that some alternative procedure was required (18,134). One of the major problems with the present system was the length of time the protease mixture was in concentrated solutions during various manipulations. The direct result of this problem, aside from the inconvenience of the length of time taken, was autolysis and decreased yields of active enzyme. A secondary problem of considerable importance was the inefficiency of the system — less than 100 mg of purified SGT was produced per preparation.

These problems were first approached by turning our attention to the dialysis procedure. If this was eliminated the time factor could be reduced considerably. A solvent system was devised which would both buffer at pH 5.0 and be volatile on lyophilization. The latter property obviated the use of dialysis in the procedure.

The second problem to be solved involved increasing production efficiency and resolution. Since the loading capacity and resolution of CM-cellulose was poor, this exchanger was abandoned and another ion exchanger investigated. SE-Sephadex C-50 had a capacity of 2.3 ± 0.3 meq/g and included the additional feature of a molecular sieving property. The buffer chosen was pyridine acetate which

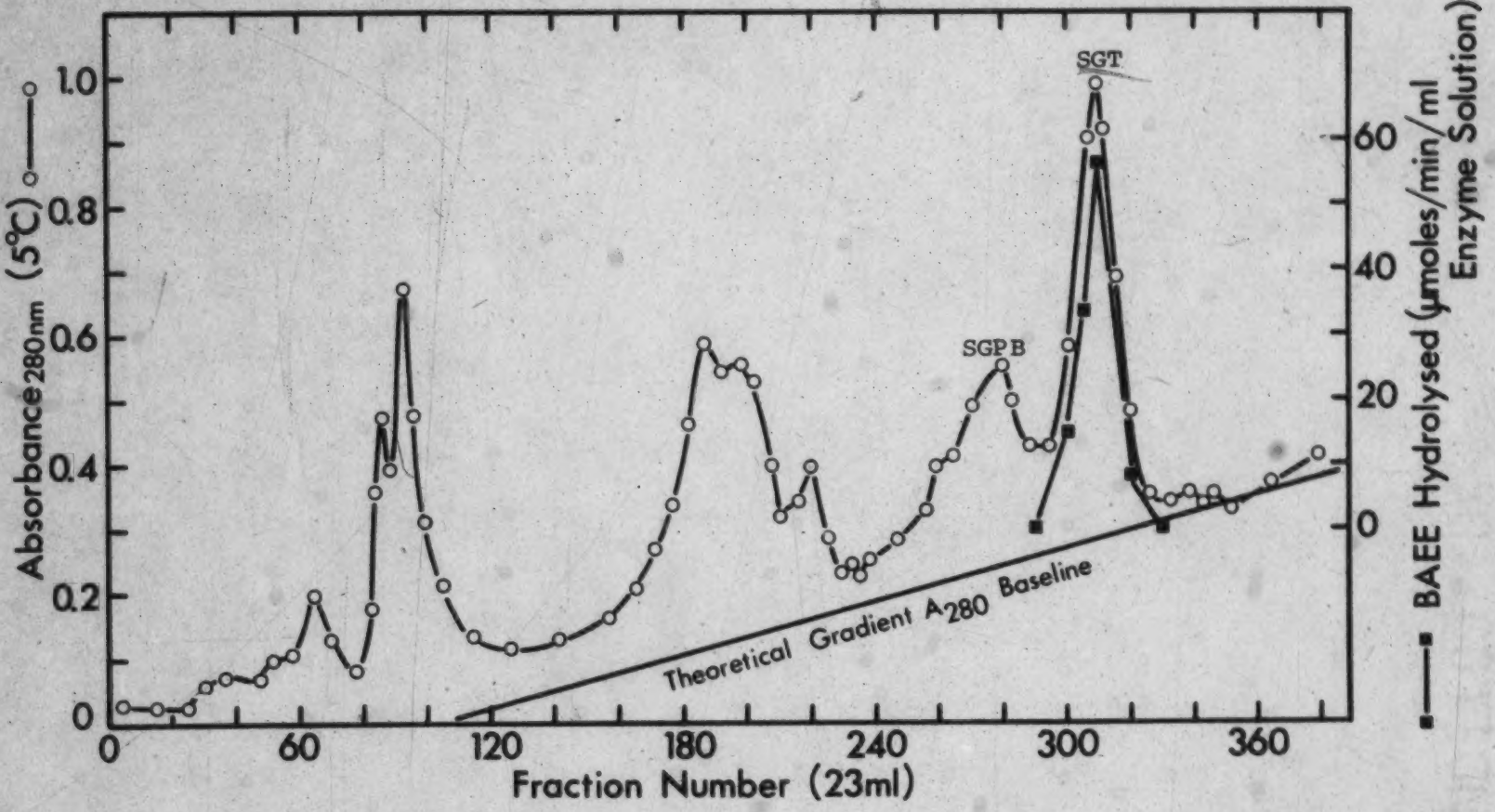


Fig. 7: SE-Sephadex chromatography of 4.5 g of crude pronase on a 5 x 100 cm column equilibrated with 20 mM pyridine-acetic acid buffer, pH 5.0. The column was developed in a linear fashion from 0.02 to 0.5 M, with respect to pyridine, using 8.6 ℓ of buffer. Fractions were collected at 100 ml/hr.

has a maximum buffering capacity at pH 5.2 (147). This buffer was used at pH 5.0 throughout our studies. Although this buffer absorbs in the ultraviolet (λ max = 255 nm), the proportion of absorbance due to maximum concentrations of buffer was not a serious drawback to reading column effluent at 280 nm. The gradient system used began at 0.02 M and increased in a linear fashion to 0.5 M using 8.6 ℓ of buffer. A typical fractionation of crude Pronase is shown in Figure 7. In this separation, a yield of between 4 and 5% purified SGT was routinely obtained. Although the quality of this SGT was acceptable for sequence studies, there was minor contamination from the preceding peak labelled SGPB (Streptomyces griseus Protease B) which catalyzed the hydrolysis of ATEE (N-acetyl-L-tyrosine ethyl ester) (20). This result precluded the use of SGT prepared by this procedure from most kinetic and physical studies where highly purified enzyme was required. Since the SGPB contamination was clearly the result of tailing, no more than 5 g of crude Pronase could be applied to the column without considerable contamination of SGT with SGPB. Despite this shortcoming the SE-Sephadex column procedure solved many of the previous problems. For example, dialysis was no longer required as the buffer was volatile on lyophilization and consequently, the time required to maintain the protease in a concentrated solution was decreased.

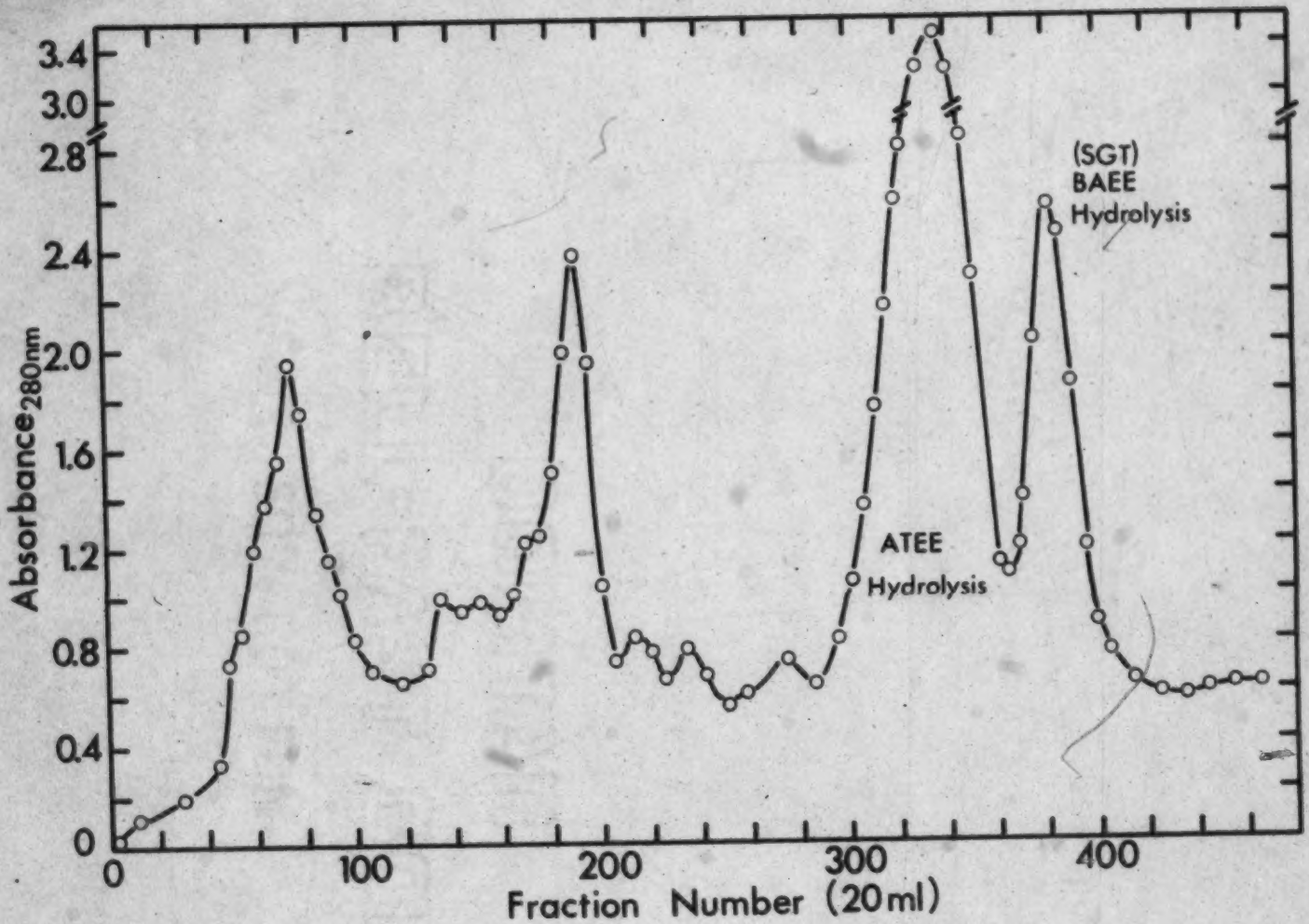

3.4
3.0
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
Absorbance280nm
0
100
200
300
400
Fraction Number (20ml)
(SGT)
BAEE
Hydrolysis
ATEE
Hydrolysis

Figure 8

Fig. 8: See Facing Page.

CM-Sephadex chromatography of 20 g of crude pronase on a 5 x 100 cm column equilibrated with 20 mM pyridine-acetic acid buffer, pH 5.0. The column was developed in a linear fashion from 0.02 to 0.75 M with respect to pyridine, using 13 liters of buffer. Fractions of 20 ml were collected at 100 ml/hr.

In addition, the yields of SGT were increased four to five times that of the previous two column system. This increase in yield was believed to be primarily a result of decreased autolysis losses.

(2) CM-Sephadex C-50 Chromatography of Pronase

While the preceding system appeared to provide adequate separation of Pronase components when samples of 5 g or less were applied to a 5 x 100 cm column, a modification of this technique using CM-Sephadex and the same buffer, fractionated as much as 20 g of crude Pronase with high resolution. The modification increased the final buffer concentration but not the slope of the gradient. This system is described in Chapter II, Methods section A1. The elution profile of a typical separation of 20 g of crude Pronase with this system is presented in Figure 8. The order of elution of enzymatic components is exactly that observed with SE-Sephadex in Figure 7. It should be noted, however, that considerable variation in quantities of various components of Pronase have been observed when different batches are compared. The apparent differences in the SE-Sephadex profile of Figure 7 and the profile in Figure 8 are due to this phenomenon. For example, the enzyme denoted SGPA (Streptomyces griseus Protease A) was designated as one of the three major endopeptidases in Pronase and reported in a publication arising, in part, out of this thesis work (20). This

enzyme has all but disappeared from subsequent batches of Pronase as is clearly denoted in Figure 8. Peak I in this figure corresponds to a pigmented material which has no known proteolytic activity while peak II corresponds to the leucine aminopeptidase activity reported by Jurasek et al (1971) (20). The SGPA activity found and reported in the previous publication is represented by the first small peak eluting after peak II in Figure 8. This was shown to be the case by demonstrating that peak II had no activity towards ATEE while the following small peak was active. The quantity of SGPA had decreased approximately 65% as estimated from the maximum absorbance at 280 nm. It should be included that, conversely, the quantities of SGPB protease, corresponding to peak III in Figure 8, had increased considerably in size over that previously published, while little or no change has been observed in peak IV, corresponding to the SGT peak.

As a result of a chemical modification procedure used in structural studies to be described later, a difference in the purification of SGT on CM- and SE-Sephadex was revealed. SGT purified only by chromatography on the CM-Sephadex system appeared to be slightly contaminated with protease(s) resistant to 8 M urea. This contaminant was apparently responsible for the degradation of DFP inhibited SGT (DIP-SGT) when chemical modification in this denaturant was attempted. (see Chapter II, Methods section K2). This problem will be

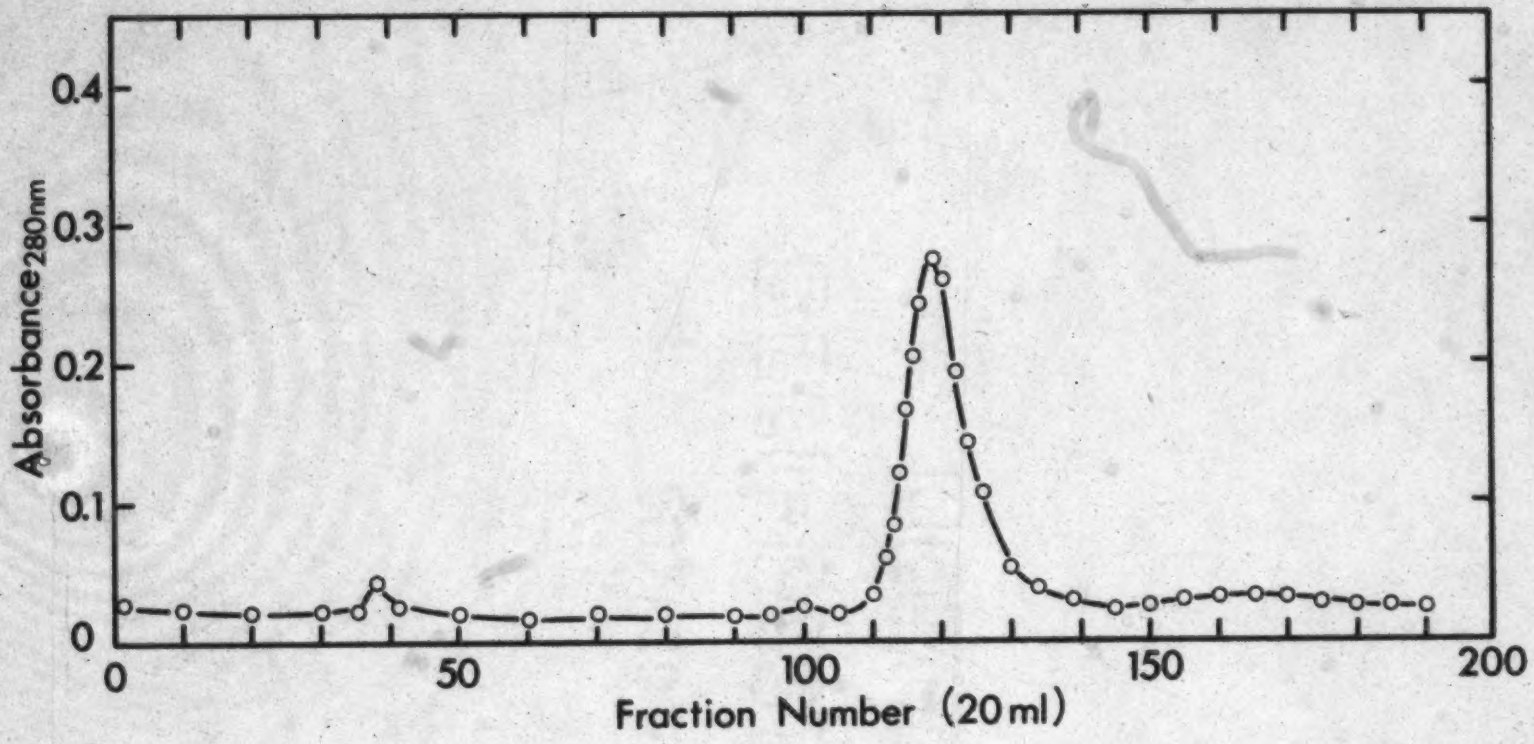


Fig. 10: Bio-Rex-70 chromatography of 100 mg of SGT previously fractionated on both CM- and SE-Sephadex. (cf. Figs. 7 and 8) The sample was applied to a 5 x 75 cm column equilibrated with 0.1 N NaOH-cacodylic acid buffer, pH 6.10 and the column developed at 60 ml/hr with the same buffer.

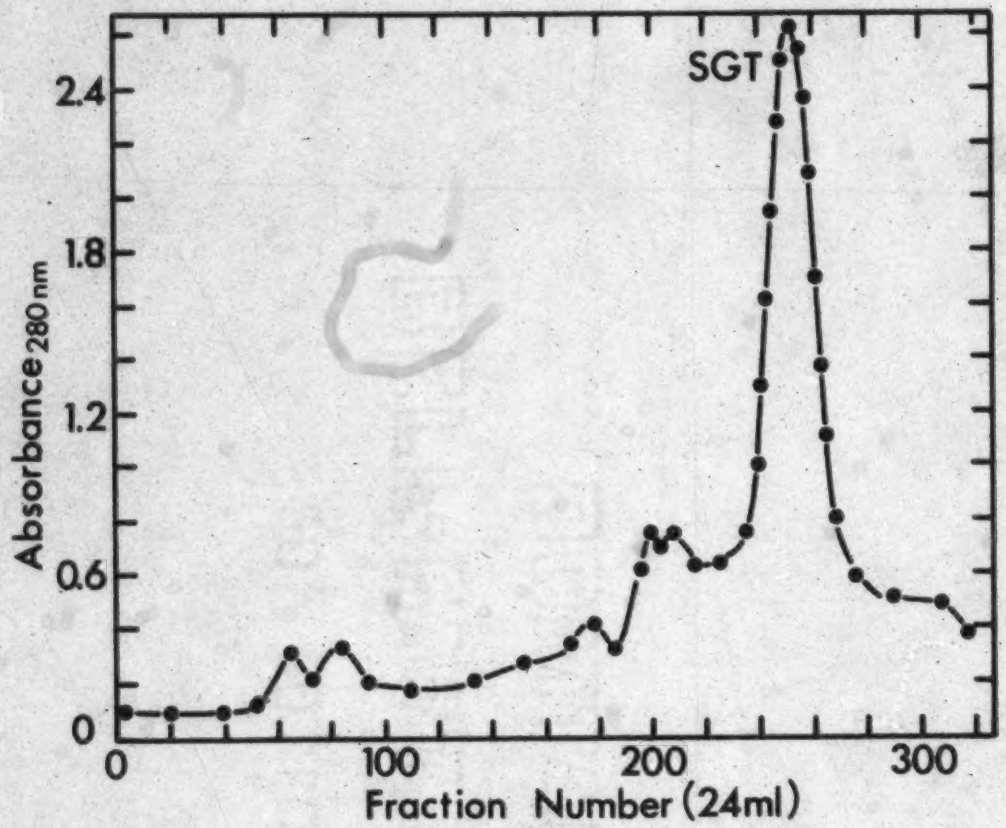
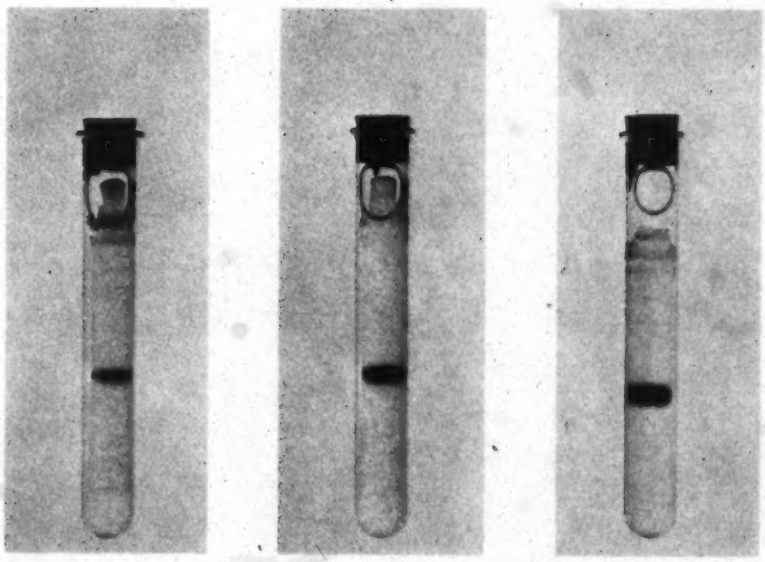


Fig. 9: SE-Sephadex chromatography of 800 mg of SGT previously fractionated on CM-Sephadex chromatography. (cf. Fig. 8) The sample was applied to a 5 x 100 cm column equilibrated with 0.1 M pyridine-acetic acid buffer, pH 5.0, and developed with a linear gradient increasing to 0.5 M with respect to pyridine. The flow rate was 100 ml/hr.

elaborated upon during the presentation of results from the structural studies to follow. However, in order to increase the purity of SGT, Pronase was routinely chromatographed on CM-Sephadex, and the SGT enzyme was further purified on SE-Sephadex C-50 using a modified shallow gradient buffer system, described in Chapter II, Methods section A1. The modified gradient was linear from 0.1 M to 0.5 M in 8.6 ℓ of pyridine acetate buffer, pH 5.0. A profile of SGT previously fractionated on CM-Sephadex and finally chromatographed on this modified SE-Sephadex system is shown in Figure 9. It is evident that SGT purified by chromatography on CM-Sephadex contains uncharacterized material which is separable on this modified gradient system. For specificity studies requiring highly purified enzyme, Bio-Rex-70 chromatography, using the sodium cacodylate, pH 6.1 system, was employed (109). This system is described in Methods, section A2. SGT purified on both CM-Sephadex and SE-Sephadex as described above was applied to a 5 x 75 cm column of Bio-Rex-70 and eluted with 0.1 N sodium hydroxide-cacodylic acid buffer. Figure 10 represents a profile of 100 mg of SGT isolated using both Sephadex systems and finally purified on Bio-Rex-70. The peak tube from this fractionation contained enzyme with a specific activity of 246 U/mg SGT ——the highest value ever attained during these experiments.



20 µg 40 µg 80 µg

Fig. 11: Polyacrylamide gel electrophoresis of 20, 40, and 80 µg samples of SGT. Electrophoresis was performed at 12.5 volts per cm for 2 hours at pH 4.3 using a 7.5% acrylamide gel.

B. Homogeneity of SGT

Streptomyces griseus trypsin, purified on CM-cellulose and Bio-Rex-70, was subjected to polyacrylamide electrophoresis at pH 4.3 using 20, 40, and 80 µg samples. It is evident from Figure 11 that the preparations were pure according to this criterion. A sample of TLCK inhibited SGT purified on the 'CM-SE Sephadex' column chromatography systems, defined earlier, was electrophoresed in sodium dodecylsulphate (SDS)-polyacrylamide gel according to Weber and Osborn (1969) (148) and was found to run as a single homogeneous band. The sample was electrophoresed in an 8% slab polyacrylamide gel, 0.1% (w/v) in SDS. (I am grateful to Dr. P. Johnson who ran this sample while engaged in his own research.)

Further indications of the general homogeneity of SGT have come from column procedures. Figure 9, presented earlier, is a profile of SGT purified on SE-Sephadex after an initial fractionation on CM-Sephadex. The highly symmetrical nature of the peak is indicative of the general homogeneity of the protein. This was corroborated by performing elution chromatography on a portion of the protein isolated from the above peak. The weak cation exchanger Bio-Rex-70 was used for this fractionation and the result was shown earlier in Figure 10. It is worth reiterating that enzyme assayed from the peak tube of this profile demonstrated a specific

activity of 246 U/mg SGT ——the highest value attained in any purification procedure. Furthermore, since the enzyme was active at pH 6.1, the pH at which this chromatography was performed, the small peak at the breakthrough volume was, in all probability, the result of an insignificant amount of autolysis. Therefore, SGT purified by the CM-Sephadex-SE-Sephadex procedure described above, was highly homogeneous by this criterion.

It should also be stated that results obtained from sequence studies indicate that the molecule is homogeneous. At no time during the isolation and characterization of peptides arising from digests of SGT, was there a peptide which could not be accounted for in the primary sequence of the molecule.

Lastly, specificity studies using the oxidized insulin A and B chains as substrates, indicated that while a minor enzymic contaminant existed in highly purified SGT, this material could be separated from the 'trypsin-like' activity after passage through a 195 cm Bio-Rex-70 column. Lyophilization of fractions which contained this contaminating activity, but not SGT, showed no evidence of proteinaceous material, indicating that the actual quantity of contaminant was very low. Furthermore, it was shown that at low enzyme to substrate molar ratios (1:250) where the effect of the spurious activity was minimized, hydrolysis of the B chain of

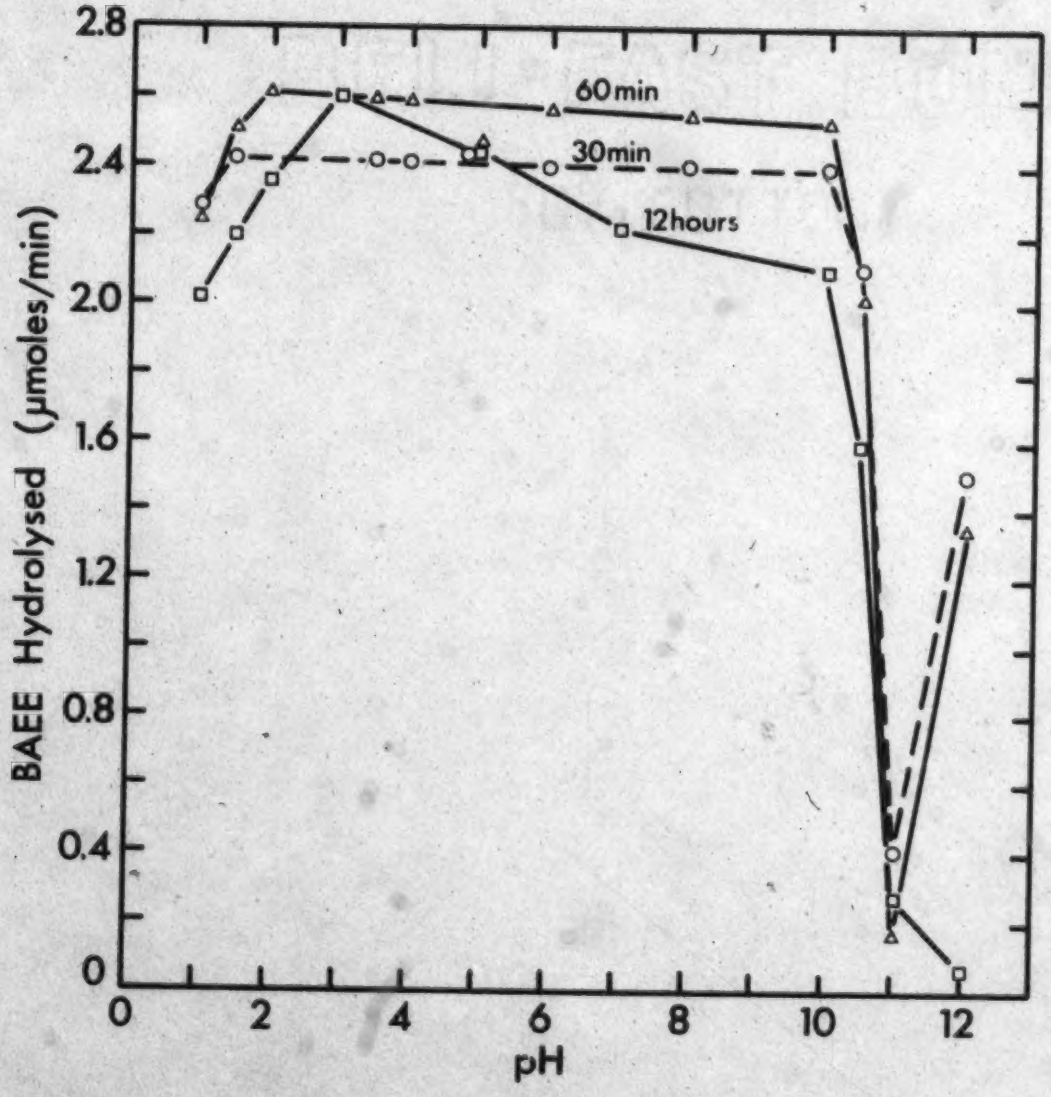


Fig. 12: Plot of SGT activity versus pH at various incubation times. The experiment was carried out at 0° in the absence of calcium ion and activities were measured spectrophotometrically at pH 8.0 using BAEE as substrate.

oxidized insulin was carried out at a rate approximately 90% greater than that of the A chain. These results indicated that SGT, as routinely purified, was highly homogeneous with the exception of a contaminant present in amounts too small to quantitate by any criterion other than the presence of trace spurious enzyme activity.

In summation, it is clear that the routine method of preparation of SGT yields a highly homogeneous enzyme which contains trace amounts of 'non-trypsin-like' activity separable by column chromatography on Bio-Rex-70.

C. Calculation of the Molar Extinction Coefficient

Using the method detailed in Chapter II, Methods section O, the molar extinction coefficient at 280 nm was calculated to be 3.96×10^4 M^{-1}/ cm. A more useful relationship for routine laboratory work, however, was $A^{1cm}_{280nm} = 1.73$ for a 1 mg/ml solution of SGT.

D. The Stability of SGT in the Presence and Absence of Ca^{2+}

The stability studies with SGT were carried out at 0° and 30°, with and without Ca^{2+} ion, for varying times and assayed spectrophotometrically as detailed in Chapter II. The stability of SGT at 0° was very good throughout the pH range 3.0 to 10.0 as shown in Figure 12. Even after 12 hrs at pH 10.0 the loss in enzyme activity

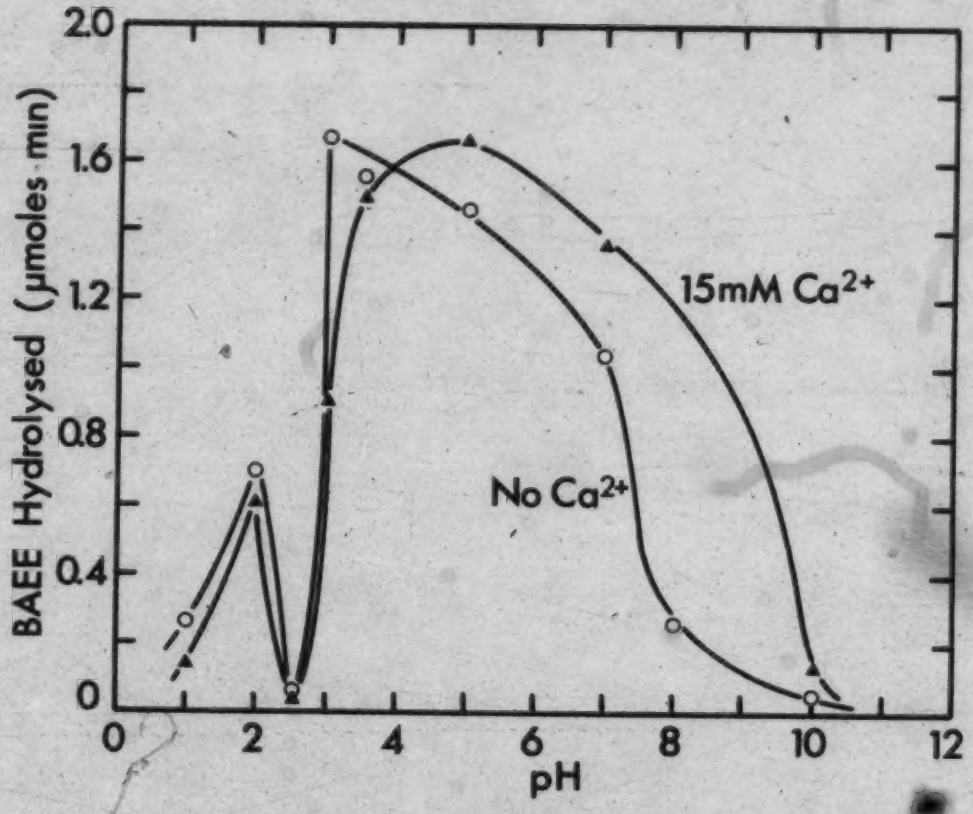


Fig. 14: Plot of SGT activity versus pH after incubation for 24 hours at 30°. The experiment was carried out in the presence and absence of 15 mM calcium ion and activities were measured spectrophotometrically at pH 8.0, using BAEE as substrate.

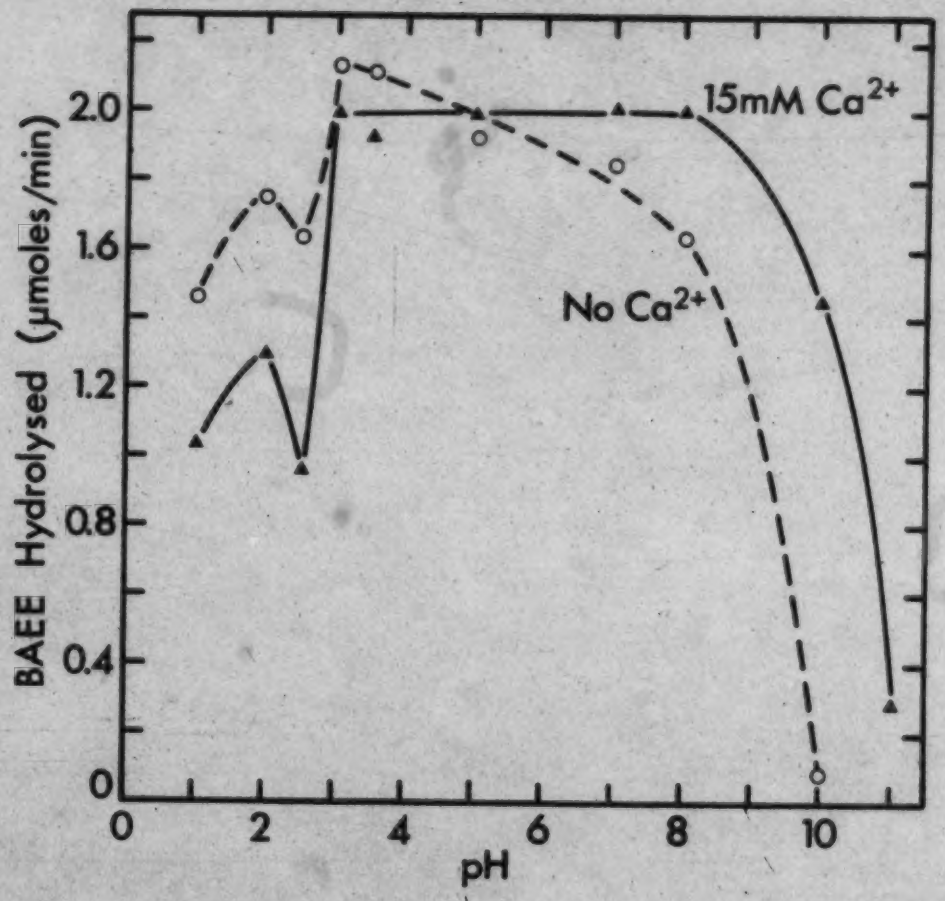


Fig. 13: Plot of SGT activity versus pH after incubation for 25 min at 30°. The experiment was carried out in the presence and absence of 15 mM calcium ion and activities were measured spectrophotometrically at pH 8.0 using BAEE as substrate.

was only 17%. However, as observed in Figure 12, activity falls off precipitously between pH 10 and pH 11.0. Apparently however, approximately 50% of the activity can be preserved up to 1 hr at pH 12.0, after which nearly all is lost at 12 hrs. This phenomenon was reproducible. The results at 30° with and without $CaCl_2$ are shown in Figures 13 and 14. It is evident that the pH optimum for extended incubations (24 hrs) at 30° in the presence of 15 mM $CaCl_2$ is pH 5.0 whereas in the absence of $CaCl_2$ the optimum is pH 3.0. Indeed, the stability of the enzyme towards higher pH values is increased in the presence of $CaCl_2$. This stabilizing effect of Ca^{2+} is observed in bovine trypsin and other serine proteases (47). It should also be pointed out that like the pH 11.0 phenomenon observed at 0° there was a pH of rapid inactivation observed at pH 2.5, at 30°. This was followed by a pH of higher stability at pH 2.0, similar to the pH 12.0 effect at 0°. Again these were highly reproducible effects.

The latter phenomenon of increased stability at pH 2.0 and 12.0 might be attributed to an equilibrium of native and denatured enzyme. For example, at pH's greater than 3.0 SGT is less stable, undergoing an increasing loss in activity as the pH increases. This is presumably due to an equilibrium between native and denatured enzyme, where the denatured form is digested. However, at

pH values greater than 11.0 the native SGT concentration may become so small that large amounts of the reversibly denatured trypsin accumulates. Similarly, at pH 2.0 SGT shows an aberrant increase in stability which might indicate a crossover point between a pH where the reversibly denatured enzyme is digested by trace amounts of native enzyme (pH 2.5), and a pH where a more drastic irreversible denaturation occurs (pH 1.0). Indirect proof that material incubated at pH 12.0 and 2.0 is in fact, reversibly denatured is obtained when an aliquot of the solution is introduced to the assay system at pH 8.0. The rate of the reaction increases with time over approximately 5 seconds, at which time it becomes linear. This phenomenon was not observed at any other pH studied.

It should be mentioned that this phenomenon of increased stability at very high and very low pH values, was also observed by Green and Neurath (1954) with bovine trypsin and essentially the same explanation for the result was given at that time (49).

A final stability study was done at a higher SGT concentration (5 mg/ml) in 50 mM sodium acetate buffer, pH 4.9 containing 0.2 M KCl at 20° for 20 hrs. The enzyme lost only 10% of its original activity over this period of incubation as assayed spectrophotometrically.

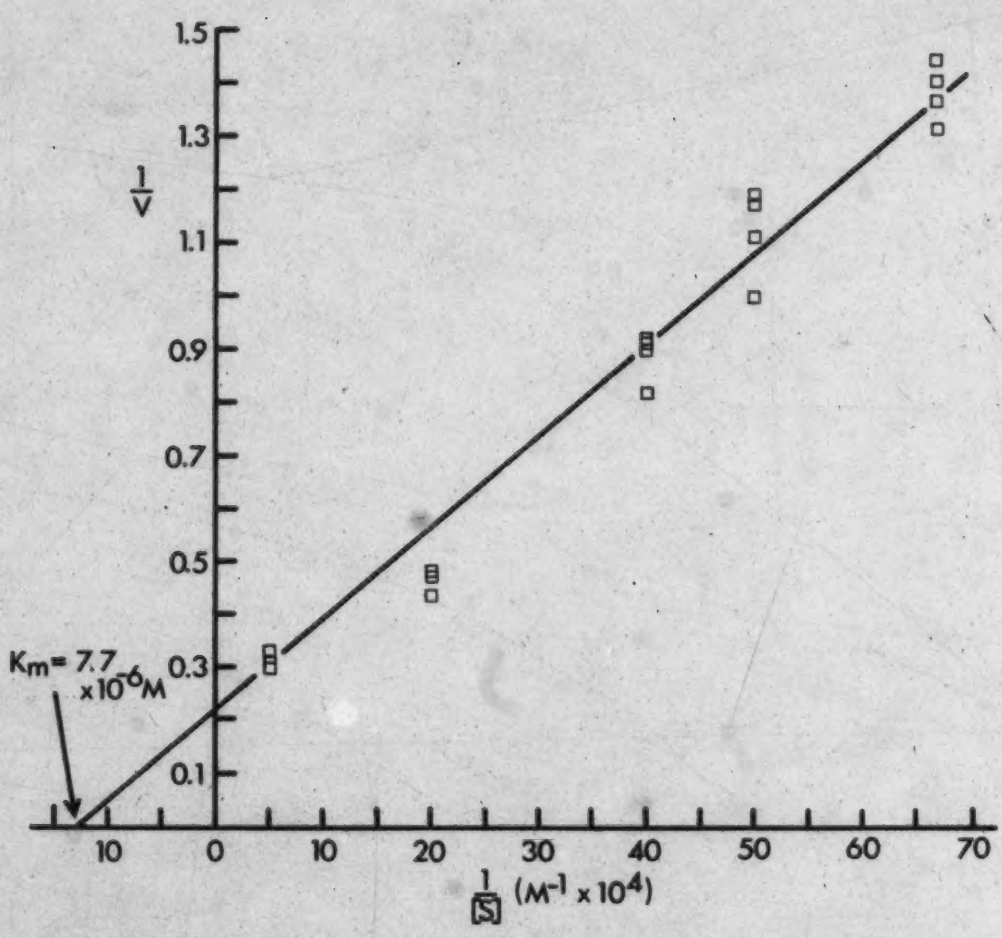


Fig. 15: Lineweaver and Burk reciprocal plot for estimation of the apparent Michaelis constant for SGT. The reaction was carried out in a nitrogen purged pH-stat, at pH 7.5, with the temperature regulated at 30°.

E. Determination of the Michaelis Constant for SGT

Reaction rates for hydrolysis of BAEE substrate were followed using the standard pH-stat assay procedure modified to contain a 1500 ml beaker in a 5 gallon circulating water bath under a nitrogen atmosphere. This modification was found necessary as SGT, like trypsin, has a K_m of approximately 10^{-5} M for ester substrates. Therefore, in order to obtain reliable initial rate values for low substrate concentrations, one liter substrate volumes were required. The stock enzyme concentration was 0.3 mg/ml (SA = 148 U/mg SGT) from which 100 µl was added to varying substrate concentrations made up in 1 ℓ volumes of the standard pH-stat assay solution. The temperature was equilibrated at 30° before each run and no CO_2 uptake was observed over a 35 min period using this system. Three to four determinations of each substrate concentration were made. K_m values, calculated according to the method of Lineweaver and Burk (1934) (see Figure 15), varied from a low of 3.5×10^{-6} to 1.5×10^{-5} M for five determinations (149). The mean was 7.7×10^{-6} M. This value is in excellent agreement with Yoshida et al (1971) (37) who found a K_m of 7.3×10^{-6} M for SGT and with Gutfreund (1965) who found a K_m of 10^{-5} M for bovine trypsin (150).

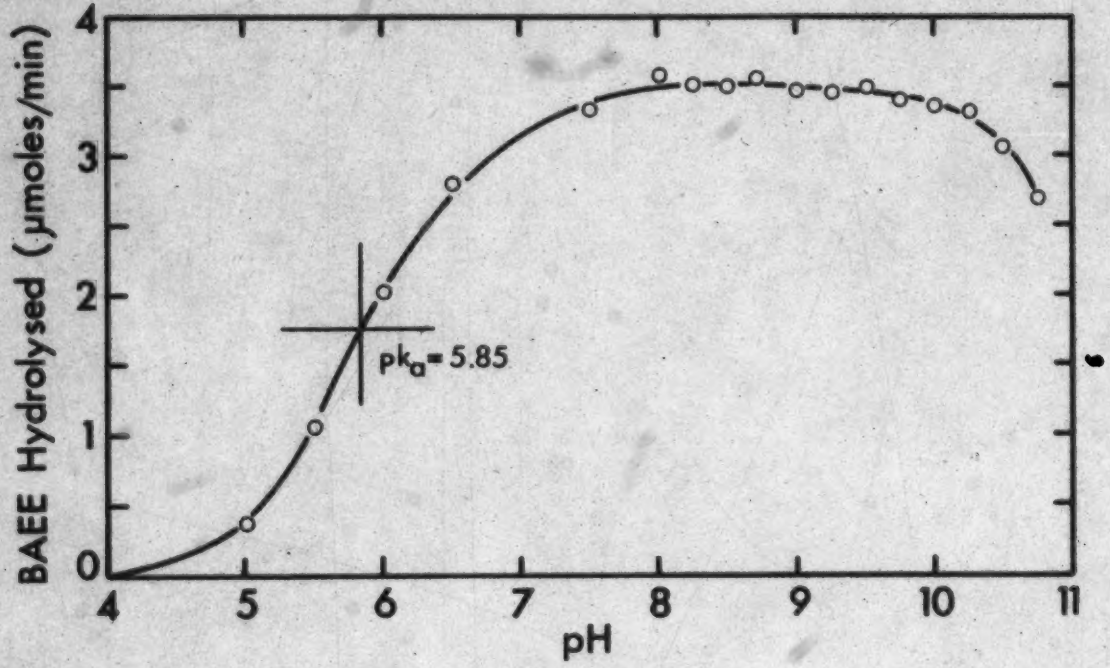
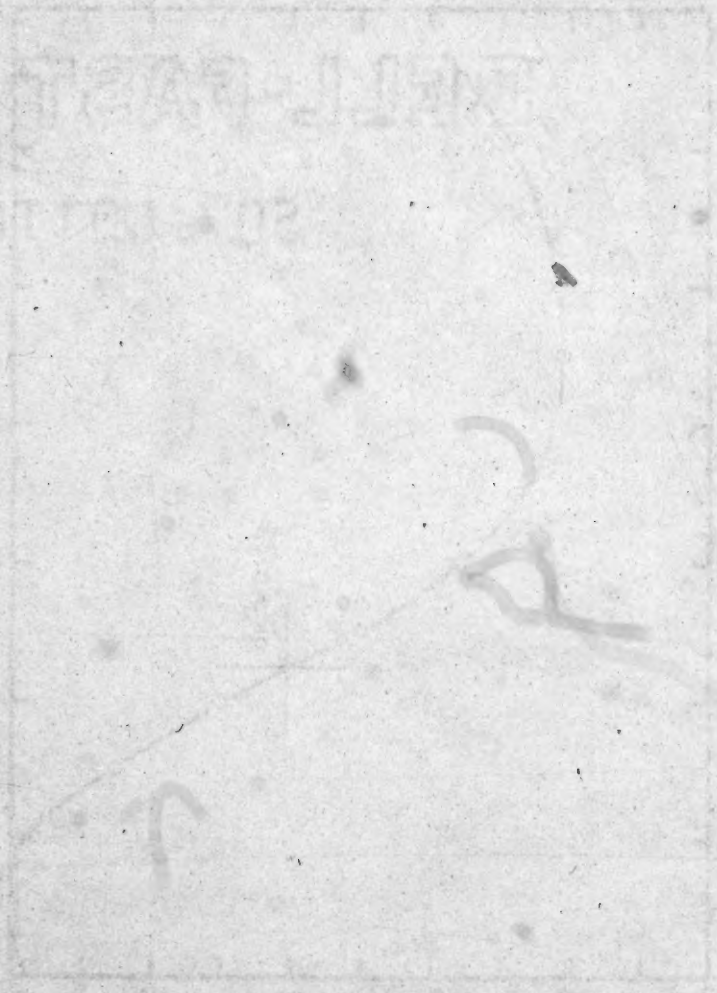


Fig. 17: Plot of the pH dependence of SGT activity as measured with a nitrogen purged pH-stat using BAEE as substrate. The intersecting lines indicate the observed pk$_a$.



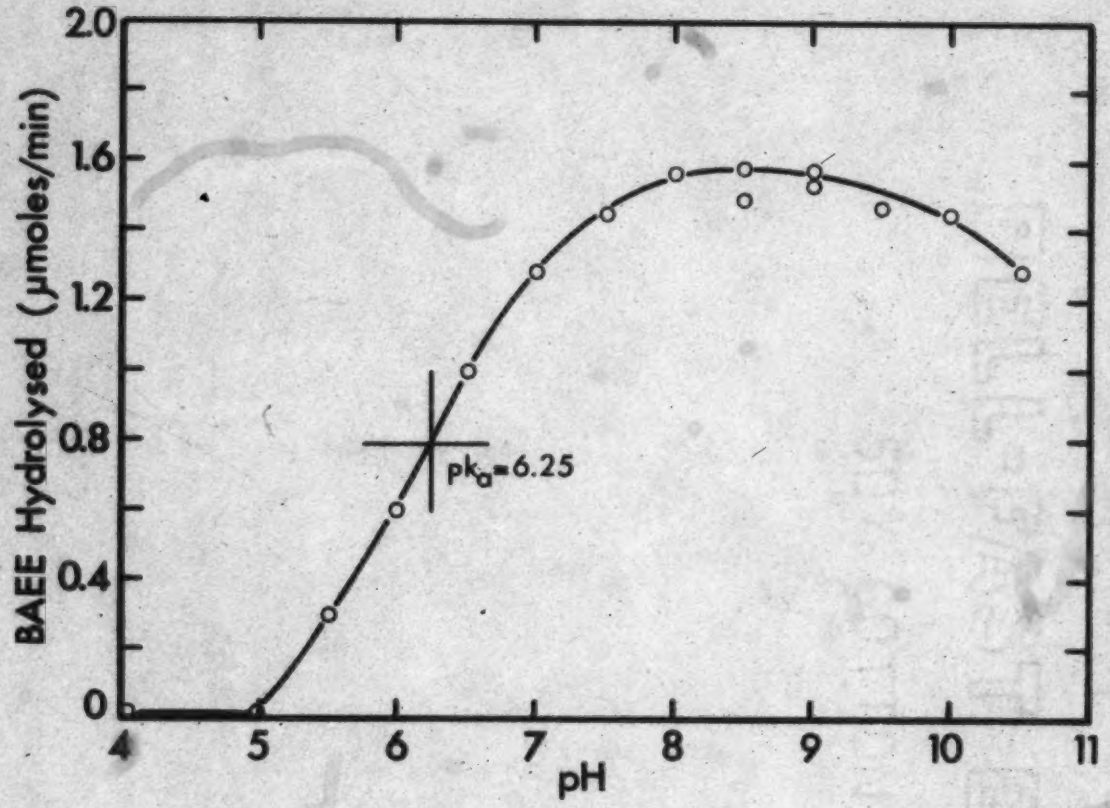


Fig. 16: Plot of the pH dependence of bovine trypsin activity as measured with a nitrogen purged pH-stat using BAEE as substrate. The intersecting lines indicate the observed pk_a.

F. Enzyme Activity Versus pH Profile for SGT

Trypsin and SGT were compared in an activity versus pH study to ascertain the approximate pK_a of the group or groups being titrated. The assays were done using the standard pH-stat procedure as described in Chapter II, Methods section I-2. The apparent pK_a for titration of trypsin was found to be 6.25 (Figure 16) which agrees precisely with the value published earlier by Gutfreund (1955) (151). On the other hand, the SGT apparent pK_a was 5.84 (Figure 17), averaged over four determinations ranging from 5.7 to 5.95. It is not unlikely that the group being titrated at approximately pH 6 is the imidazole side chain on the active histidine-57 which should have a pK_a of approximately 6.0 (152). Since a charged histidine-57 would be unable to function in the hypothesized charged transfer complex which requires that the histidine relay a proton, the enzyme would be inactive at pH values below approximately 5.0 (24). This is in agreement with the data presented here for SGT.

It is also evident from Figure 17 that the activity falls off again at approximately pH 10.0. It might be pointed out that a major difference between the activity profile observed for SGT (Figure 17) and that observed for bovine trypsin (Figure 16), is the relative invariance of SGT activity between pH 8.0 and 10.0, whereas the trypsin activity falls off earlier, after pH 9.0.

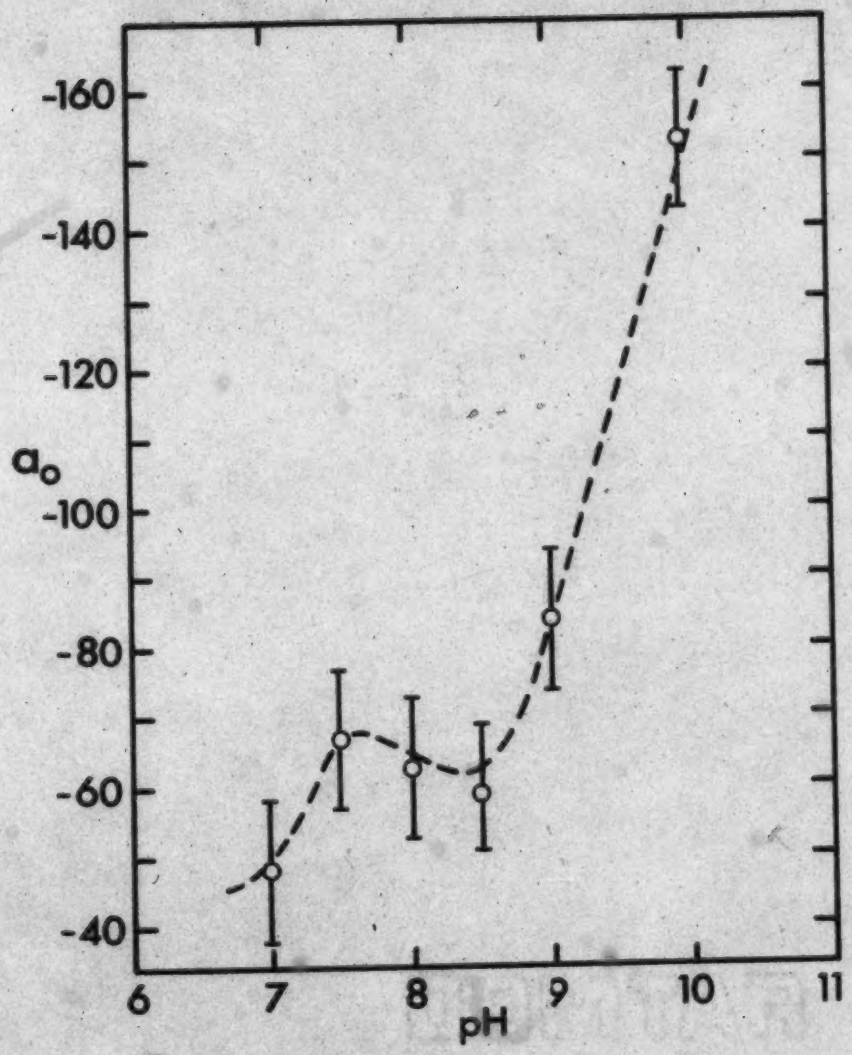


Fig. 18: Plot of the Moffit equation constant a_o versus pH, as calculated from ORD measurements in the wavelength range 550 to 300 nm. The expected error in estimation of a_o, according to McCubbin *et al* (1966), is applied to each point (153).

TABLE VIII

MOFFIT-YANG PARAMETERS AND α-HELIX ESTIMATIONS AS A FUNCTION OF pH

pH	a_o	b_o	% Helix
7.0	-48	-130	20
7.5	-67	-125	20
8.0	-63	-123	19
8.5	-59	-129	20
9.0	-84	-121	19
9.9	-163	-122	19

G. Optical Rotatory Dispersion and Circular Dichroism Studies

In view of the present interest in the possible importance of the NH_2-terminal α-amino group, as discussed in the introduction, a group of spectropolarimetric experiments were designed to investigate whether SGT showed a pH dependent conformational change as demonstrated in δ-chymotrypsin (56) and α-chymotrypsin (69). ORD and CD experiments were performed at 10° over a pH range of 5.3 to 10.5. In the case of ORD measurements in the wavelength range 550 to 300 nm, the data were collected at several pH values from pH 7.0 to 9.9 and the parameters a_o and b_o calculated according to the equation of Moffit and Yang (127). These values, together with estimations of percentage α-helix are presented in Table VIII. It should be noted that the b_o and α-helix data indicate no change in helical content nor extensive unfolding of the enzyme between pH 7.0 and 9.9.

A plot of a_o versus pH is shown in Figure 18, as this is the only value which appeared to vary as a function of pH. The expected error in estimation of a_o according to McCubbin *et al* (1966) is applied to each point and a best fit drawn (153). Since the value a_o is a complex parameter dependent upon environmental interactions with the surface side chains as well as internal side chain interactions, the sharp decrease in

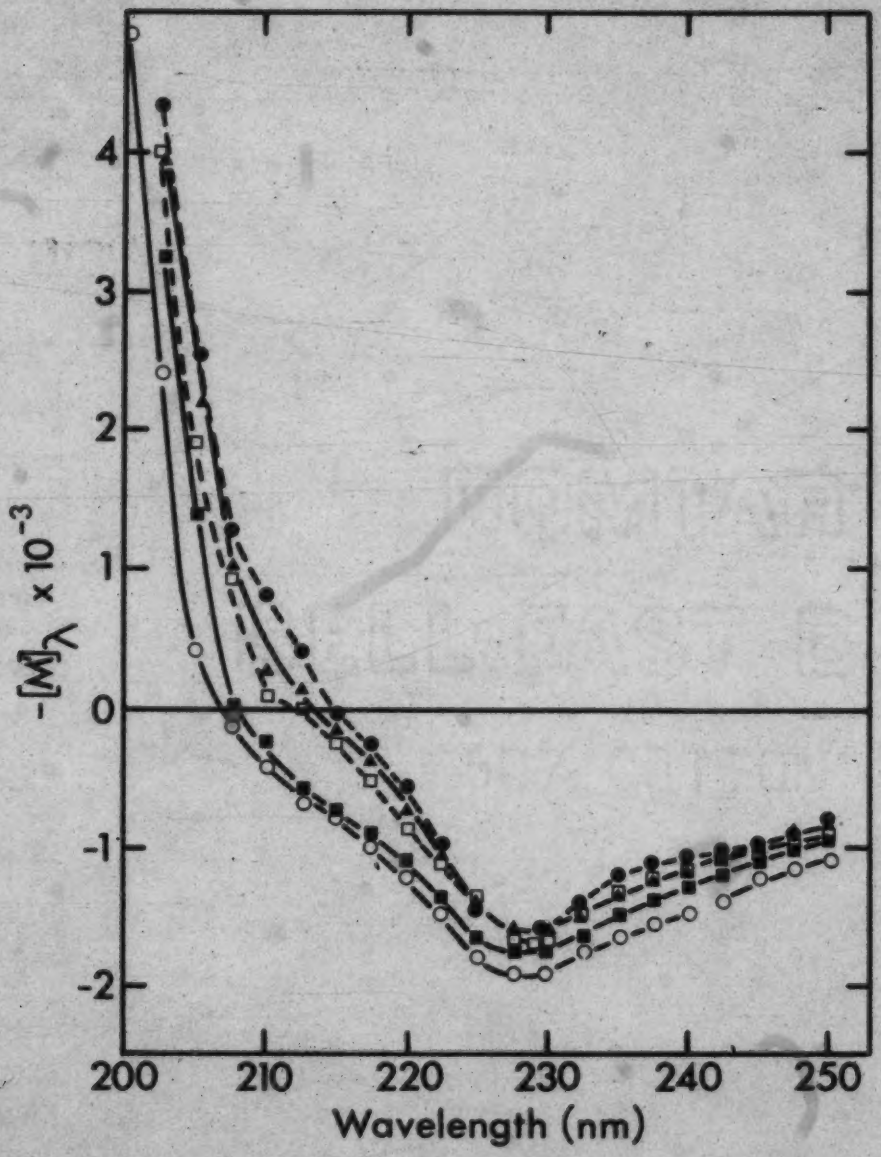


Fig. 21: Optical rotatory dispersion spectra of SGT performed over the pH range 7.3 to 10.1. ●—●, pH 7.3; □—□, pH 8.3; ○—○, pH 8.6; ▲—▲, pH 8.9; ■—■, pH 10.1.

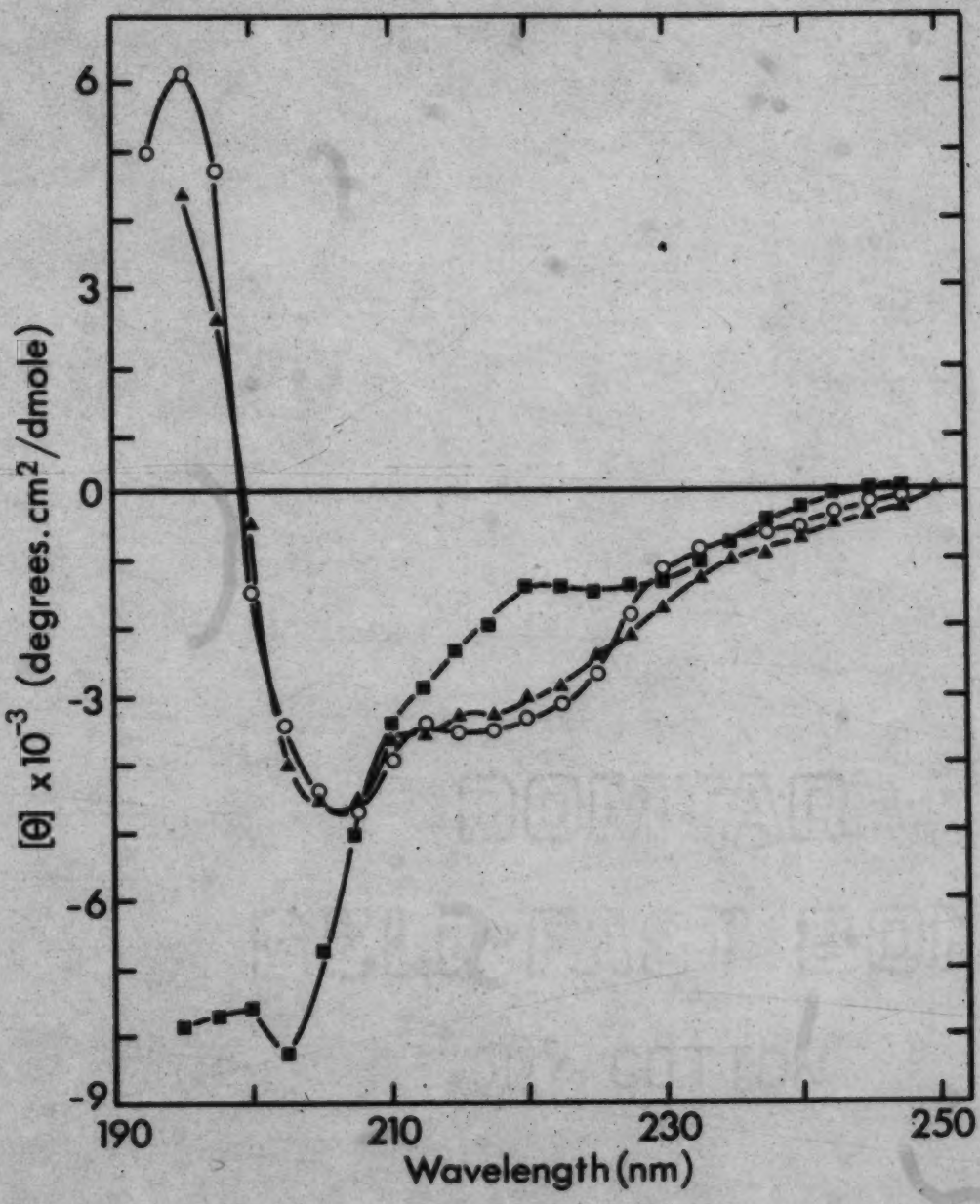


Fig. 20: Circular dichroism spectra of SGT performed over the pH range 8.5 to 10.5. o——o ,pH 8.5; ▲——▲ ,pH 9.3; ■——■ ,pH 10.5.

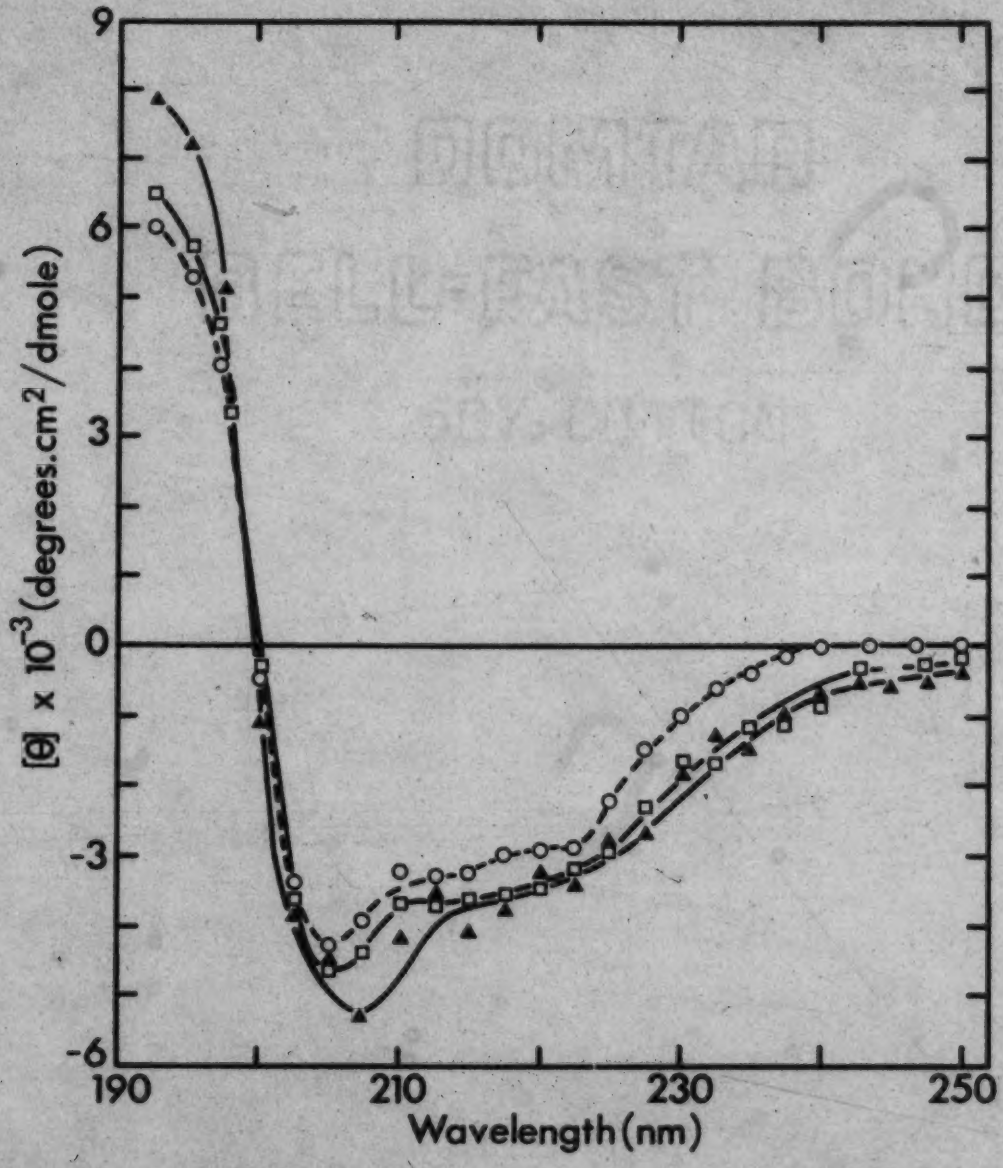


Fig. 19: Circular dichroism spectra of SGT performed over the pH range 5.3 to 7.6. ▲——▲, pH 5.3; ○——○, pH 6.5; □——□ , pH 7.6.

the value of a_o from pH 9.0 to 10.0 might well be interpreted as demonstrating the unfolding of SGT (154).

That these results further indicate that this is not a titration of an acidic group is demonstrated by CD studies done in the far ultraviolet between the wavelengths 250 and 190 nm. These scans are depicted in Figures 19 and 20, which show an insignificant change in enzyme conformation between pH 5.3 and 9.3 followed by a large transition at pH 10.5 which is clearly a denaturation of the enzyme. The same experiment using ORD in the ultraviolet range, between the wavelengths 250 and 200 nm, showed essentially no major conformational changes between pH 7.3 and 10.1. This result is depicted in Figure 21.

These results are in agreement with those found using α-lytic protease from Myxobacter 495 and indicate that the drop in activity observed at pH values greater than 10.0 as shown in Figure 17, is associated with a general denaturation and unfolding of the enzyme (75). There is no evidence from these results to implicate the role of an ionized α-amino group in the activity and stabilization of the enzyme. Indeed, if such a group is involved these data indicate that it only becomes available for titration on general unfolding and denaturation of the enzyme at pH values greater than 10.0.

H. Attempts to Block the NH_2-Terminal α-Amino Group by Acetylation and Carbamylation

In conjunction with the experiments described previously, it was of considerable importance that an attempt be made to block the NH_2-terminal α-amino group and measure the resultant enzyme activity. Since there is no known zymogen for SGT ε-amino groups on lysine residues could not be blocked prior to activation and liberation of the α-amino group as was done with trypsin (79) and chymotrypsin (56). An attempt was made to acetylate the native enzyme directly, as accomplished by Kaplan and Whitaker (1969) (76) with α-lytic protease, and if activity was not reduced after reaction at the NH_2-terminal, a conclusion might be drawn. If any other result arose, no conclusion could be drawn.

Acetylation was carried out as described in Chapter II, Methods section K5. The protein precipitated even before the reaction, performed under the conditions as described by Openheimer et al (1966) (56), was completed. In spite of this awkward result, it was decided to attempt to assay the soluble fraction and perform an NH_2-terminal analysis by using a modified dansyl procedure in which the reagent solution contained 8 M urea. The specific activity of the control unreacted enzyme was 157 U/mg SGT, while the soluble reacted enzyme had a SA = 59.3 U/mg SGT. Furthermore, the soluble material

had all ε-amino groups acetylated, as estimated by dansylation, but the DNS-Val (dansyl-valine) and more exactly DNS-Val-Val (dansyl valylvaline) dipeptide was found on hydrolysis of the enzyme. (The NH_2-terminal sequence of the molecule is Val-Val-Gly-Gly-etc.) These data indicate that, although complete reaction of lysine residues was effected, reaction at the NH_2-terminal α-amino group was either incomplete or absent altogether. This latter conclusion is in agreement with the work of Scrimger and Hofmann (1967) (78) and Robinson et al (1973) (79) who showed that denaturation was necessary for chemical modification of the NH_2-terminal α-amino group in bovine trypsin. The latter group was able to circumvent the problem of autolysis losses during NH_2-terminal α-amino modification by first guanidinating the zymogen, thus blocking the ε-NH_2-lysine groups without loss of charge. The ε-guanidinated trypsinogen was then activated at pH 3.5 by an acid protease and shown to be active. This modification of trypsinogen decreased the sites of autolysis to only the available arginine residues on the molecule. ε-guanidinated trypsin was then denatured in 6 M guanidinium/HCl at pH 3.0 and carbamylated with KNCO at pH 8.0. The uncarbamylated enzyme demonstrated 90-95% of the original activity on renaturation. Since reversible denaturation techniques had not been worked out for SGT and since there was no doubt that denaturation of

SGT at pH 8.0 would result in massive autolysis, it was decided that this approach was unworkable. In order to show that this result was not specific for the acetic anhydride reagent, carbamylation was attempted according to the procedure of Rimon and Perlmann (1968) as described in Chapter II, Methods section K6 (140). Since potassium cyanate reacts 100 times faster with α-amino groups than ε-amino groups at pH 7.0, this reagent was considered a desirable alternative to acetic anhydride. In this procedure the protein stayed soluble throughout the reaction and subsequent dialysis, indicating that acetylation produced precipitation by removing cationic groups from SGT. When the control and reacted protein were again analyzed as above, the control specific activity was found to be 172 U/mg SGT while the reacted enzyme had a specific activity of 61.4 U/mg SGT. The modified dansyl NH_2-terminal procedure, containing 8 M urea, indicated that unlike the acetylation procedure, ε-amino groups were present, but in agreement with the previous reaction, the α-amino group of the NH_2-terminal was still available for dansylation. DNS-Val and DNS-Val-Val were present on hydrolysis of the dansylated protein.

It is difficult to draw any firm conclusions from these results except to say that in agreement with the previously mentioned authors, this trypsin also appears to have a very unreactive NH_2-terminal α-amino

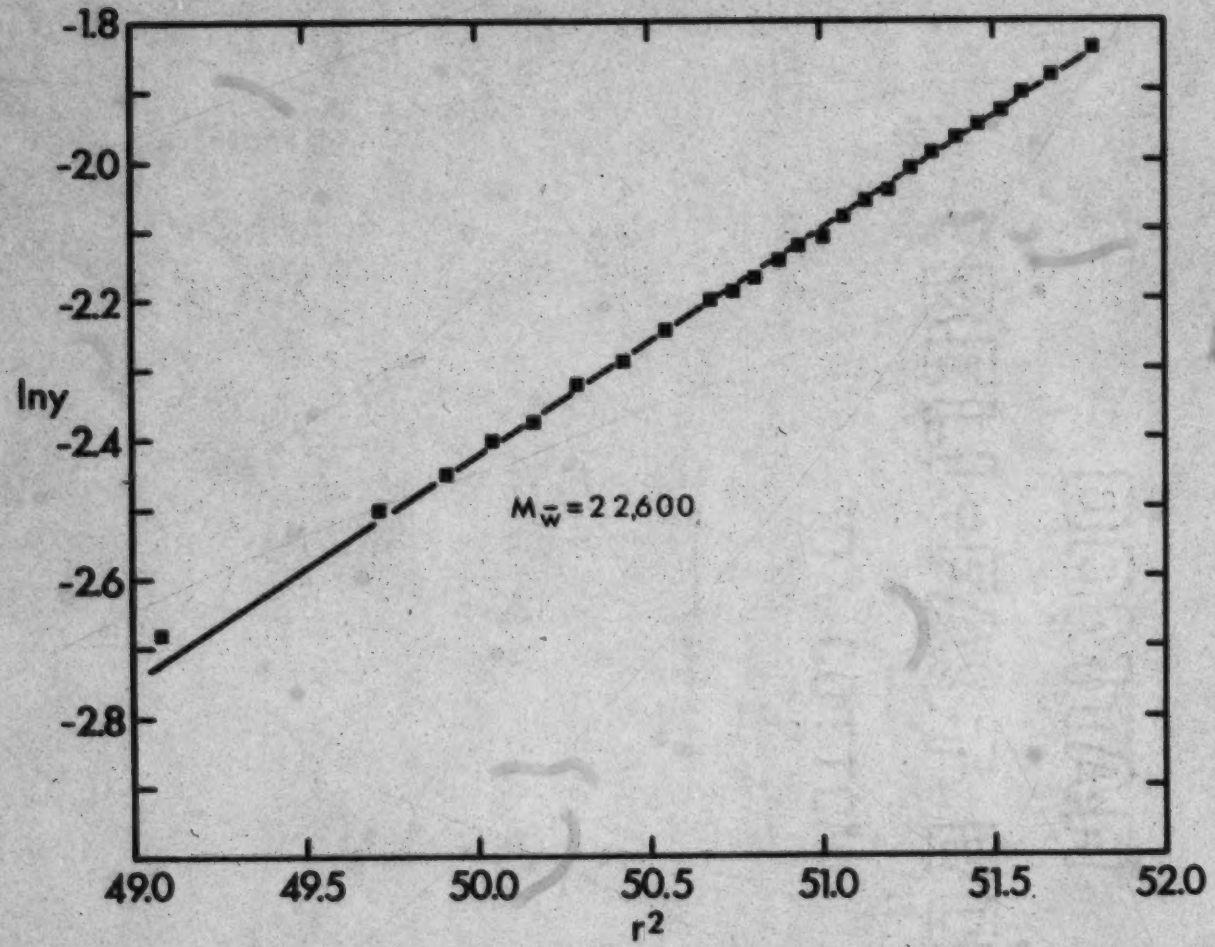


Fig. 22: Plot of the natural logarithm of the concentration (fringe displacement) as a function of the square of the distance from the axis of rotation. A low speed sedimentation equilibrium run using 1.0 mg/ml SGT in 0.5 M KCl, 50 mM sodium acetate buffer pH 5.0, at an equilibrium speed of 15.000 RPM.

group. This result had also been found with porcine elastase (34). It is tempting to suggest that these data indicate that the NH_2-terminus in this enzyme is buried in a similar fashion to those of other serine proteases such as porcine elastase (34).

I. Molecular Weight Determination By Low Speed Sedimentation Equilibrium

The molecular weight of SGT was estimated by the use of the low speed sedimentation equilibrium technique as described in Chapter II, Methods section G. Samples of 1.0 mg/ml, 1.5 mg/ml, and 3.0 mg/ml were run at an equilibrium speed of 15,000 RPM in 0.5 M KCl, 50 mM in pH 5.0 sodium acetate buffer at 20°. The partial specific volume ($\bar{v}$)was assumed to be equal to 0.717, the value calculated by Travis and Roberts (1969) for human trypsin (51). The data for the lowest concentration, that expected to behave in the most ideal manner,are shown as a ln y versus r^2 plot in Figure 22 where the $M\bar{w}_{app}$ is approximately 22,600. This is in good agreement with the minimum molecular weight of 22,918 calculated from the assumed primary sequence and well within the error of the technique. Table IX indicates the variation of molecular weight with concentration, presumably due to non-ideal behaviour as the protein concentration is increased.

TABLE IX

Run #	Protein Concentration	$\bar{M}w_{app}$
1	1.0 mg/ml	23,060 ± 690
2	1.0 mg/ml	22,600 ± 690
3	1.5 mg/ml	21,800 ± 660
4	3.0 mg/ml	20,200 ± 610

The value obtained from sequence data and ultracentrifugal analysis for the minimum molecular weight of SGT, agrees well with values quoted for other trypsins in Table IV in the introduction.

J. Studies Relating to the Substrate Specificity of SGT

Trypsin catalyzes the hydrolysis of an ester, amide or peptide bond involving the carboxyl group of a lysine or arginine residue. Jurasek, Fackre and Smillie (1969) showed that highly purified SGT (1:200 molar ratio) would not hydrolyze the oxidized insulin A chain, but would hydrolyze specifically the single arginine and lysine residues in the oxidized insulin B chain (109). Only the purest fractions from a preparation of SGT purified on CM-cellulose and twice on Bio-Rex-70 were used in the above experiments.

These results have now been expanded upon in order

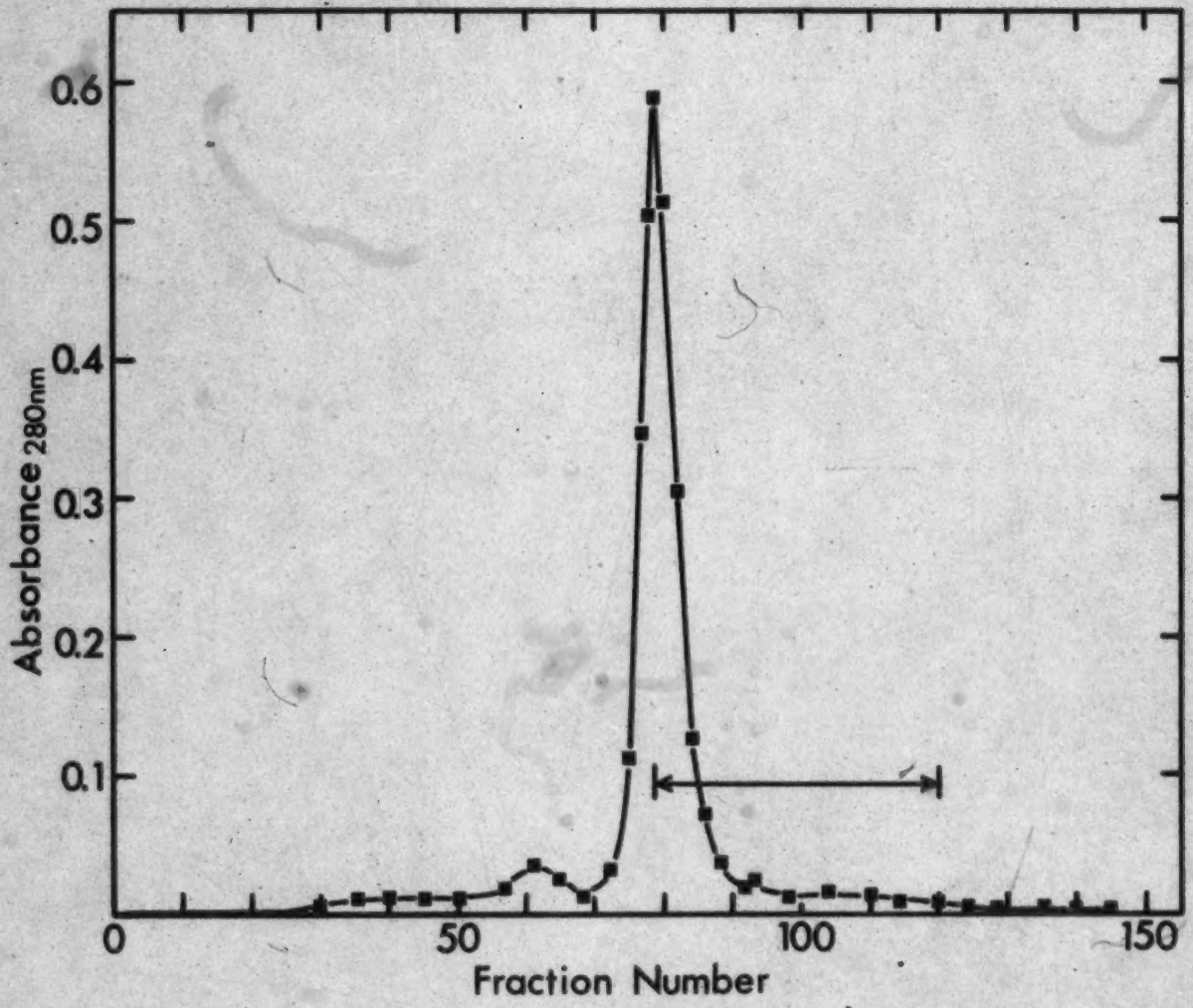

0.6
0.5
0.4
0.3
0.2
0.1
Absorbance 280nm
0
50
100
150
Fraction Number

Figure 23

Fig. 23: See Facing Page.

Bio-Rex-70 chromatography of 75 mg of SGT previously purified on SE-Sephadex. The sample was applied in 5 ml to a 2.5 x 159 cm column equilibrated with 0.1 N NaOH-cacodylic acid buffer, pH 6.10 and eluted at 18 ml/hr. The arrow designates the position of oxidized insulin A chain hydrolyzing activity, as measured qualitatively by high voltage electrophoresis and cadmium-ninhydrin staining of digests performed using samples taken across the peak.

TABLE XIV

POSITION OF 'NON-TRYPTIC' HYDROLYSIS OF OXIDIZED INSULIN A CHAIN BY SGT CONTAMINANTS

(1:1000 MOLAR RATIO)

Gly-Ile-Val-Glu-Gln-Cya-Cya-Ala-Ser-Val-Cya-Ser-Leu-Tyr-Gln-Leu-Glu-Asn-Tyr-Cya-Asn

Not isolated — A2-3

POSITIONS OF HYDROLYSIS OF OXIDIZED INSULIN B CHAIN BY SGT

(1:1000 MOLAR RATIO)

Phe-Val-Asn-Gln-His-Cya-Gly-Ser-His-Leu-Val-Glu-Ala-Leu-Tyr-Leu-Val-Cya-Gly-Glu-Arg-

B6

Gly-Phe-Phe-Tyr-Thr-Pro-Lys-Ala

B3

B4

B2

TABLE XIII

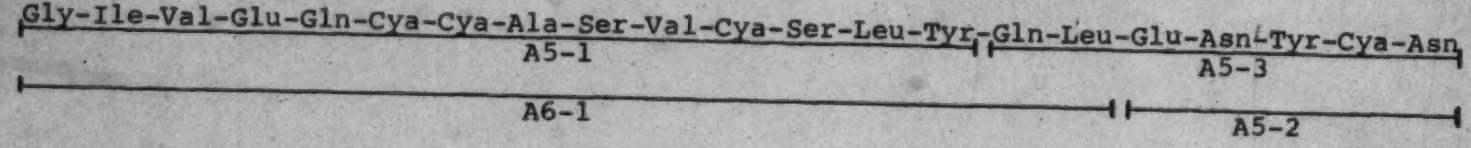
MAJOR PEPTIDES RESULTANT FROM DIGESTION OF OXIDIZED INSULIN A AND B CHAINS AT 1:1000 MOLAR RATIO (SGT TO SUBSTRATE)

Peptide	Amino Acid Composition*	Percentage Yield
A2-3	Gln,Leu,Glu,Asn,Tyr,Cya,Asn 1.00 1.12 1.00 0.82 0.64 1.25 0.82	4.1%
B2	Gly,Phe,Phe,Tyr,Thr,Pro,Lys 0.98 1.08 1.08 0.62 0.91 1.14 0.98	8.0%
B3	Gly,Phe,Phe,Tyr,Thr,Pro,Lys,Ala 0.90 1.15 1.15 0.69 0.90 1.00 0.91 0.99	1.0%
B4	Ala 1.00	1.7%
B6	Phe,Val,Asn,Gln,His,Leu,Cya,Gly,Ser,His,Leu,Val,Glu,Ala,Leu,Tyr, 1.06 0.85 0.96 0.95 0.95 0.89 0.83 0.92 0.82 0.95 0.89 0.85 0.95 1.16 0.89 0.45 Leu,Val,Cya,Gly,Glu,Arg 0.89 0.85 0.83 0.92 0.95 0.79	1.7%

* Residue values are expressed as molar ratios.

TABLE XII

POSITIONS OF HYDROLYSIS OF OXIDIZED INSULIN A CHAIN BY SGT CONTAMINANT (1:20 MOLAR RATIO)



POSITIONS OF HYDROLYSIS OF OXIDIZED INSULIN B CHAIN BY SGT AND CONTAMINANT (1:20 MOLAR RATIO)

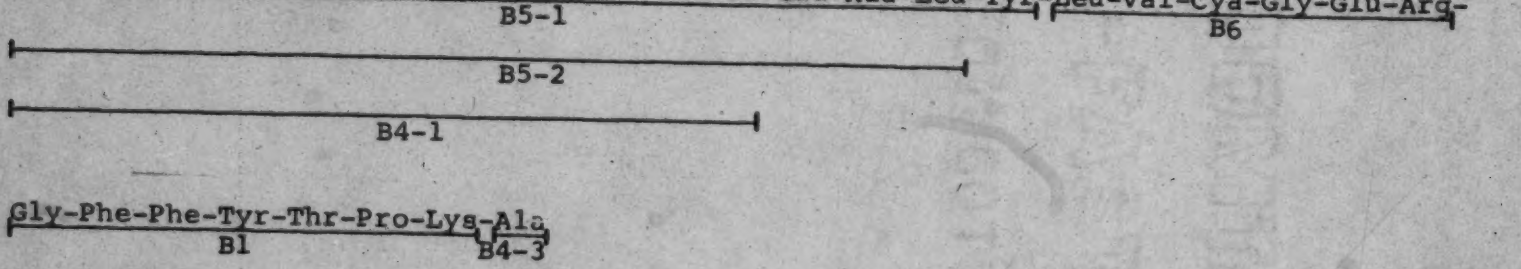


TABLE XI

MAJOR PEPTIDES RESULTANT FROM DIGESTION OF OXIDIZED INSULIN B CHAIN AT 1:20 MOLAR RATIO (SGT TO SUBSTRATE)

Peptide	Amino Acid Composition*	Percentage Yield
B1	Gly, Phe, Phe, Tyr, Thr, Pro, Lys 0.85 1.11 1.11 0.82 1.05 1.11 0.95	9.3%
B4-1	Phe, Val, Asn, Gln, His, Leu, Cya, Gly, Ser, His, Leu, Val 1.09 0.93 0.98 1.07 0.95 1.05 0.99 1.13 0.87 0.95 1.05 0.93	4.5%
B4-3	Ala 1.00	14.2%
B5-1	Phe, Val, Asn, Gln, His, Leu, Cya, Gly, Ser, His, Leu, Val, Glu, Ala, Leu, Tyr 0.87 0.98 0.98 1.02 0.95 1.04 1.03 1.15 0.97 0.95 1.04 0.98 1.02 0.96 1.04 0.63	1.6%
B5-2	Phe, Val, Asn, Gln, His, Leu, Cya, Gly, Ser, His, Leu, Val, Glu, Ala, Leu 0.85 0.95 1.06 0.99 0.95 1.04 1.02 1.14 1.08 0.95 1.04 0.95 0.99 0.92 1.04	1.2%
B6	Leu, Val, Cya, Gly, Glu, Arg 1.13 0.93 ND 1.03 1.08 0.84	10.0%

* Residue values are expressed as molar ratios.

ND = not analyzed

TABLE X

MAJOR PEPTIDES RESULTANT FROM DIGESTION OF OXIDIZED INSULIN A CHAIN AT 1:20 MOLAR RATIO (SGT TO SUBSTRATE)

Peptide	Amino Acid Composition*	Percentage Yield
A5-1	Gly, Ile, Val, Glu, Gln, Cya, Cya, Ala, Ser, Val, Cya, Ser, Leu, Tyr 1.03 0.73 0.88 1.16 1.16 ND ND 1.14 0.91 0.88 ND 0.91 1.21 0.95	11.1%
A5-2	Glu, Asn, Tyr, Cya, Asn 1.10 0.99 0.85 1.16 0.99	1.1%
A5-3	Gln, Leu, Glu, Asn, Tyr, Cya, Asn 0.96 1.06 0.96 1.03 0.88 1.07 1.03	16.2%
A6-1	Gly, Ile, Val, Glu, Gln, Cya, Cya, Ala, Ser, Val, Cya, Ser, Leu, Tyr, Gln, Leu 1.03 0.85 1.03 1.07 1.07 1.20 1.20 1.24 1.09 1.03 1.20 1.09 0.83 0.98 1.07 0.83	4.6%

* Residue values are expressed as molar ratios.

ND = not analyzed

that a very active but minor contaminant be characterized. An enzyme to substrate molar ratio of 1:20 and 1:1000 was used, as described in Chapter II, Methods section M. Although the enzyme had been purified using CM-Sephadex, SE-Sephadex and Bio-Rex-70 chromatography, it was found impossible to completely resolve the previously mentioned contaminating enzyme. On using a 1:20 molar ratio of enzyme to substrate the contaminating enzyme hydrolyzed both the oxidized A and B chains of insulin at valines, tyrosines and leucines. The resultant peptides are shown in Tables X and XI along with Table XII which shows the position of these peptides in the insulin chains. At lower concentrations of enzyme (1:1000 molar ratio) the effect of the contaminant is confined to a single hydrolysis of the A chain and a purely tryptic hydrolysis of the B chain as observed in Table XIII and in Table XIV which show the hydrolysis positions in the insulin chains. Figure 23 shows an elution profile of SE-Sephadex purified SGT chromatographed on a 2.5 x 159 cm Bio-Rex-70 column using the pH 6.1 sodium cacodylate system described earlier. Samples were taken across the peak and added to oxidized insulin A chain, on an equal absorption basis, to make a final enzyme to substrate molar ratio of 1:250. The presence of insulin A chain hydrolysis, as estimated by high voltage electrophoresis and cadmium-ninhydrin staining, is indicated on the profile. This

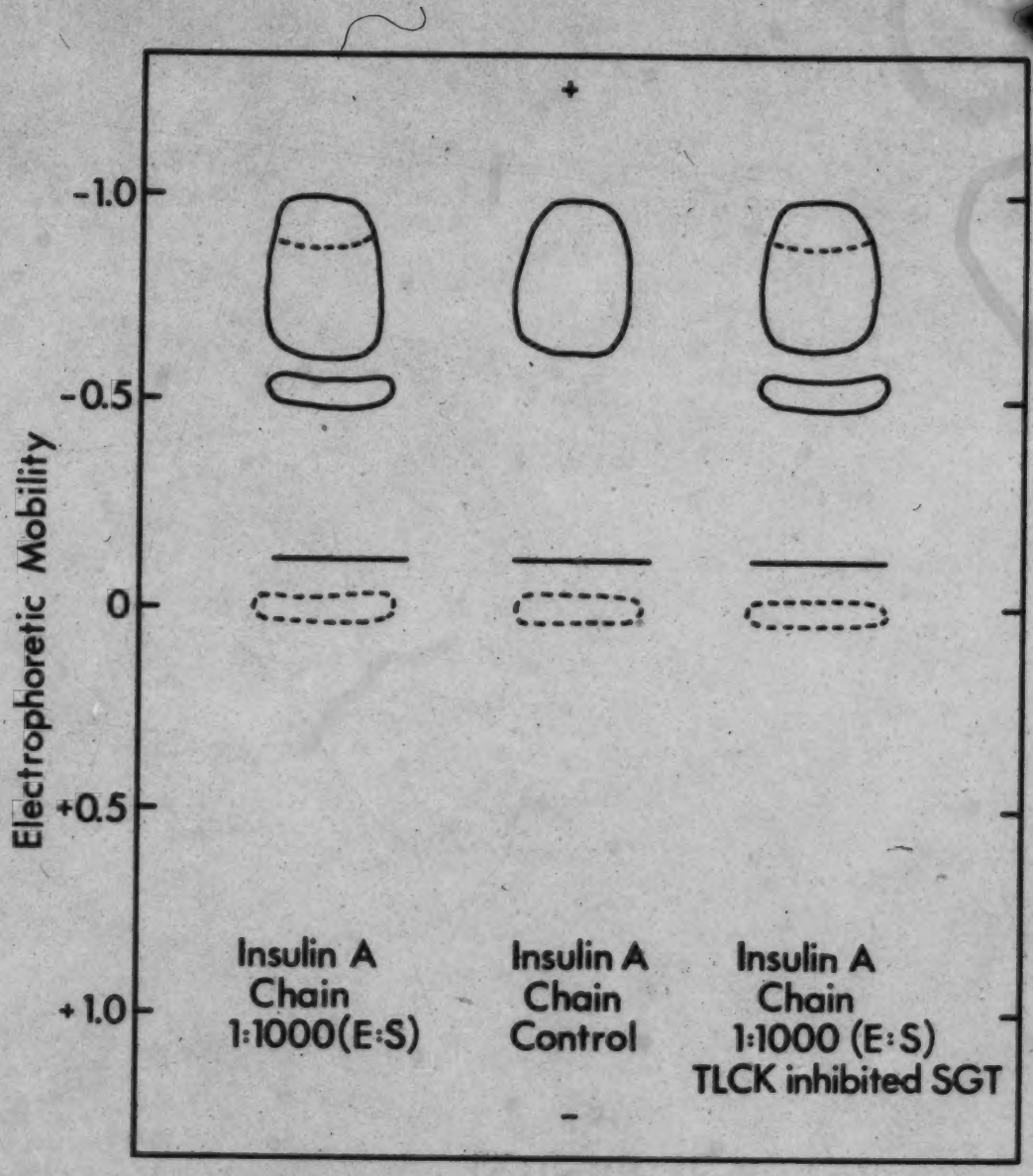


Fig. 25: Electropherogram of SGT and TLCK inhibited SGT digests of oxidized insulin A chain (1:1000 molar ratio). Electrophoresis at pH 6.5 was performed at 60 volts/cm for 45 min.

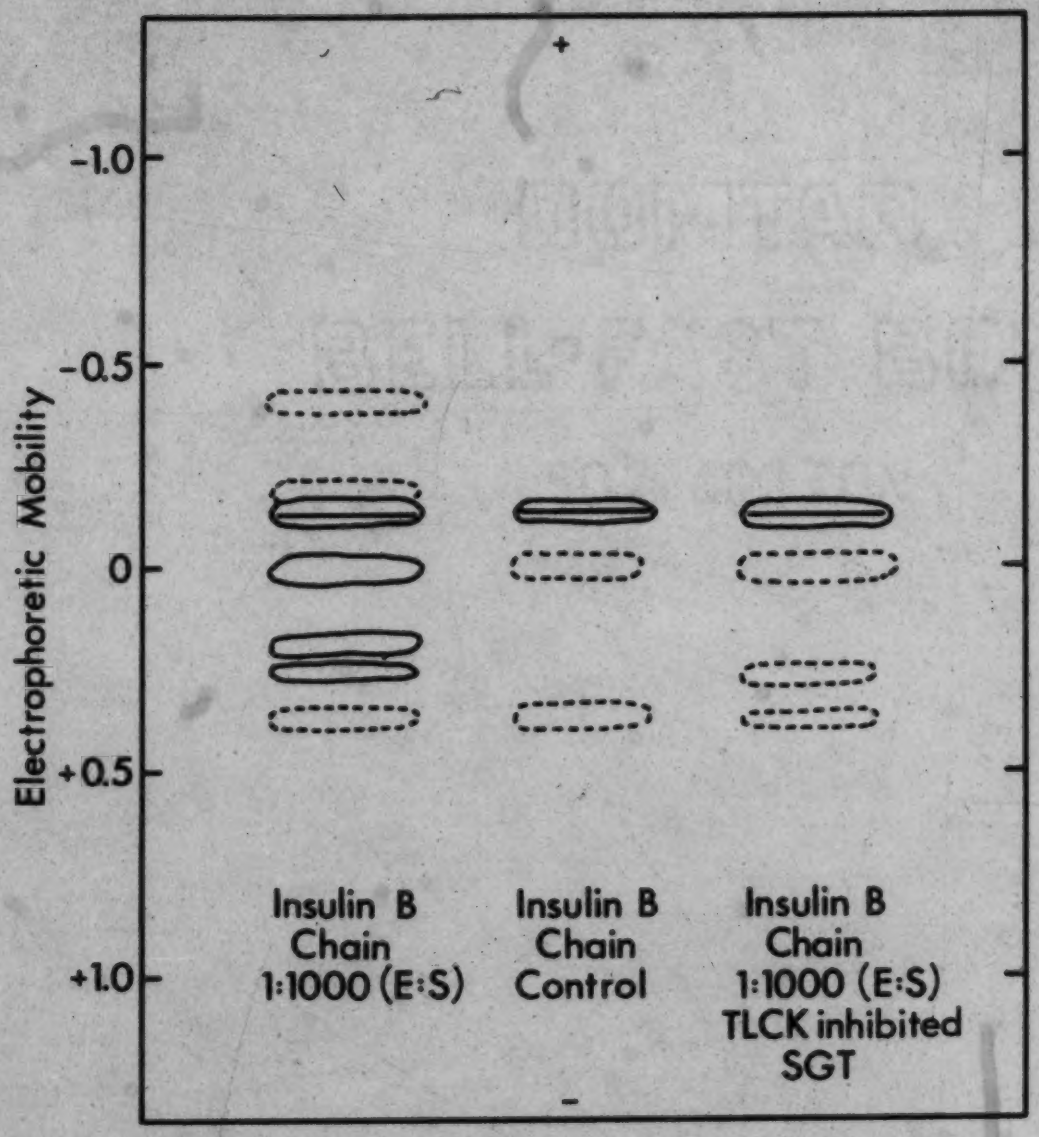


Fig. 24: Electropherogram of SGT and TLCK inhibited SGT digests of oxidized insulin B chain (1:1000 molar ratio). Electrophoresis at pH 6.5 was performed at 60 volts/cm for 45 min.

'non-tryptic' activity is observed to be partially resolved from the SGT peak and is active long after the absorbance due to SGT has reached base line. It is clear from this result that the spurious activity is separable from SGT.

Further evidence of the separate identity of this contaminating enzyme is shown in Figures 24 and 25 showing the peptides resultant from digesting oxidized insulin A and B chains at a 1:1000 molar ratio (enzyme to substrate) with SGT and with TLCK inhibited SGT. It is clear that the hydrolysis of B chain is inhibited while the hydrolysis of A chain is completely unretarded by digestion with TLCK inhibited SGT. Finally, it might be stated that this non-specific enzyme is not inhibited by TPCK, in agreement with the specificity demonstrated towards oxidized insulin B chain —in no case was a phenylalanine bond hydrolyzed. It is interesting to speculate that L-(1-tosylamido-2-tyrosyl) chloromethyl ketone might be an inhibitor of this enzyme.

In an attempt to quantitate the 'non-tryptic' activity of the SGT preparation, both oxidized insulin B chain and A chain were digested in a pH-stat as described in Chapter II, Methods section I-3, using SGT with a specific activity of 143 U/mg enzyme. Using a 1:250 molar ratio (enzyme to substrate) at pH 9.0, the B chain was digested at a rate of 230 nanomoles/min/mg SGT, while A chain was digested at a rate of 28 nanomoles/min/mg SGT.

These rates are only relative measures as the pK_a's of the nascent α-amino groups can only be approximated and therefore, the number of peptide bonds hydrolyzed cannot be calculated exactly. This relatively high rate of A chain hydrolysis is surprising however, since lyophilization of active fractions after the SGT absorbance had come to base line in Figure 23, rendered an immeasurable trace of proteinaceous material in the flask. This enzyme appears to be very active indeed. It is clear however, that at low molar ratios where the effect of the spurious activity is minimized, hydrolysis of the B chain is carried out at a rate approximately 90% greater than that of the A chain. These data, together with the inhibition studies and known separability of the two activities, give strong evidence to support the claim that SGT has a narrow specificity identical with bovine trypsin.

SGT is also known to hydrolyze many synthetic substrates in common with bovine trypsin, such as BAEE (18), Benzoyl-L-arginine p-nitroanilide (37), and N-α-Tosyl-L-arginine methyl ester (155). Trop and Birk (1970) have done extensive synthetic peptide specificity studies which essentially indicated that the enzyme hydrolyzed polylysine very well, and various neutral dipeptides to a negligible extent (156).

Finally, since a 'carboxypeptidase-like' activity

was discovered eluting after SGT on the CM-Sephadex column during preliminary investigations of the products of the preparative procedure, it was desirable to show that the SGT activity was distinct from this activity. A carboxypeptidase-A substrate, which was readily hydrolyzed by this enzyme was used (CbzGly-Leu) as described in Chapter II, Methods. The result of the experiment was that no carboxypeptidase activity could be detected when samples were taken across the SGT peak in Figure 23.

K. Inhibition Studies with TLCK and PAB — Isolation of A Modified Histidine-57 Peptide

Active-site-directed irreversible inhibitors have been instrumental in elucidating the importance of histidine-57 in maintenance of catalytic activity in α-chymotrypsin (157). In a similar manner the active-site-directed inhibitor, TLCK (L-(1-tosylamido-2-lysyl) ethyl chloromethyl ketone), being not unlike a substrate, binds to the active site and specifically alkylates histidine-57 of trypsin, further increasing the credibility of the general mechanistic similarity of catalytic action between α-chymotrypsin and trypsin (31).

In order to determine whether SGT was also irreversibly inhibited by TLCK, a time course study was devised using a seven fold molar excess of inhibitor

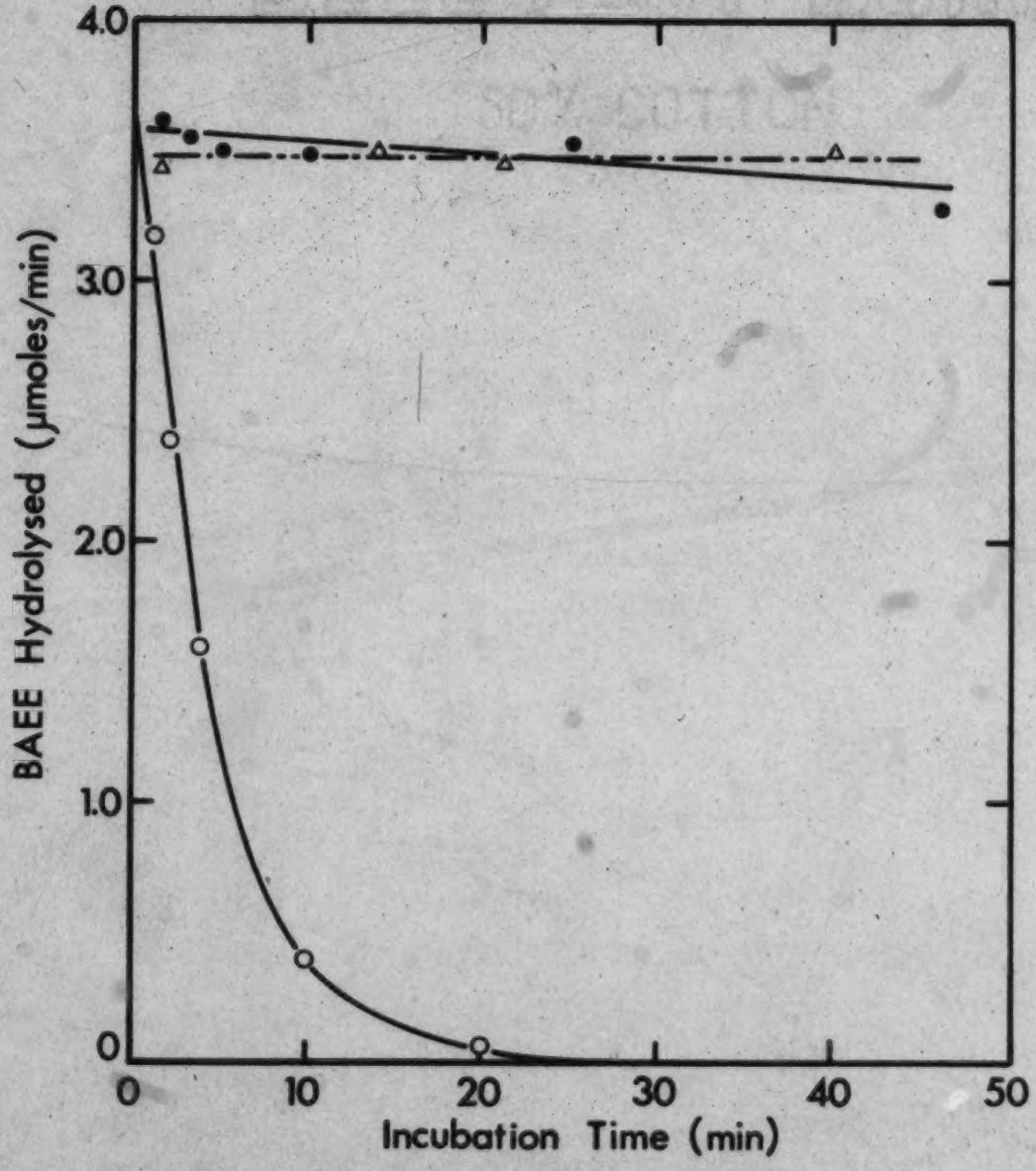


Fig. 27: Time course of inhibition of SGT with TLCK in the presence and absence of a 13 fold molar excess of p-aminobenzamidine competitive inhibitor. O——O , TLCK; ●——● , TLCK + PAB; Δ——Δ , SGT control.

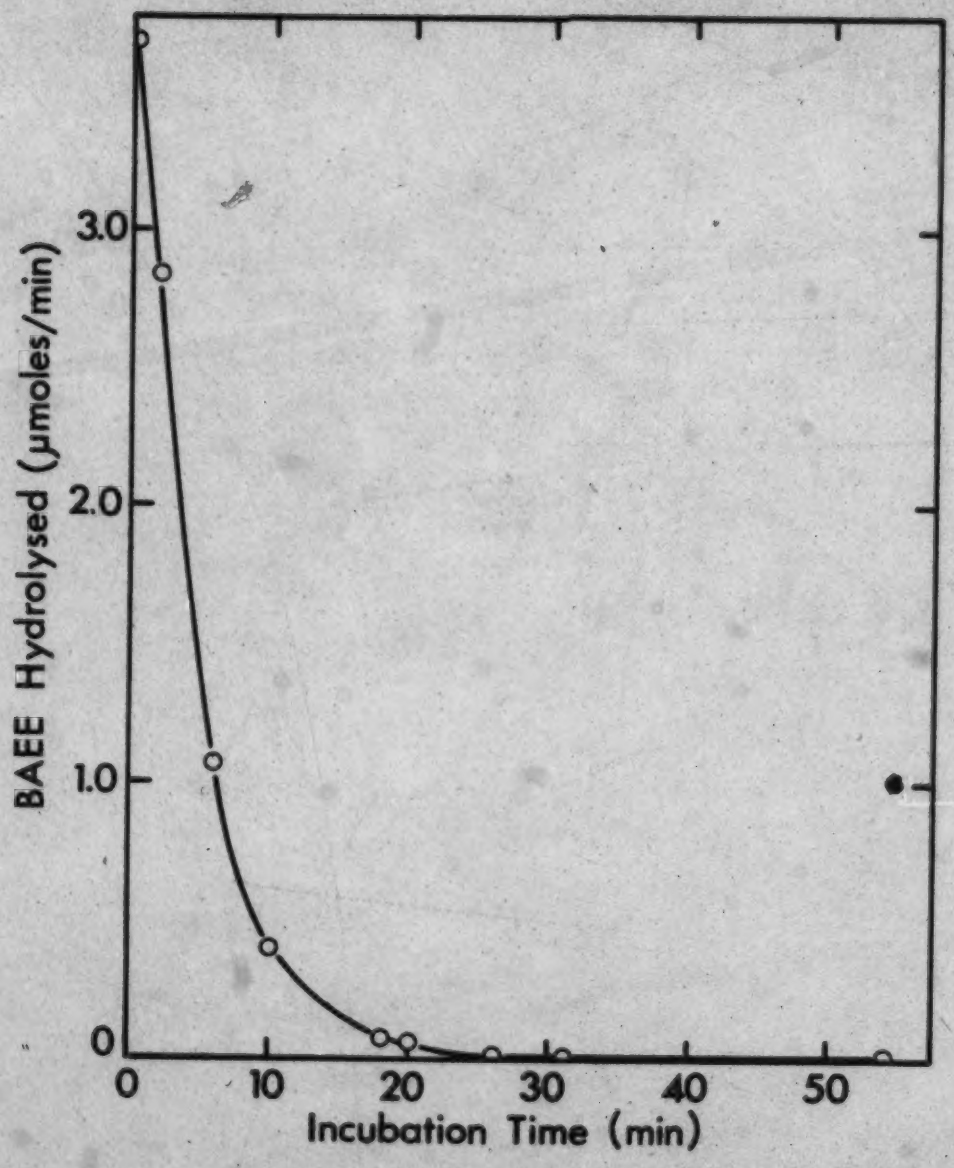


Fig. 26: Time course of inhibition of SGT with TLCK. Enzyme concentration was 2.5 mg/ml. (7 fold molar excess of TLCK) Enzyme activity was measured spectrophotometrically using BAEE as substrate.

over enzyme, as in the system described in Chapter II, Methods section K8. Activity was followed with time of incubation with TLCK and the result plotted as shown in Figure 26. Essentially, all measureable activity was lost after 25 minutes of reaction. This is in agreement with the recently published results of Yoshida et al (1971) (37).

Since TLCK is considered an active-site-directed inhibitor it should be competitively inhibited by a large excess of a reversible inhibitor such as p-amino-benzamidine. The results of the same time course previously mentioned, in the presence of a 13 fold molar excess of PAB, are shown in Figure 27. Indeed, this competitive inhibitor of trypsin (158) effectively blocks the specific irreversible inhibition of TLCK, giving strong evidence of the similarity of the active sites in bovine trypsin and SGT.

Since the only histidine in SGT is at position-57 (α-chymotrypsin numbering system) it was important to demonstrate that reaction had occurred at this residue. Conveniently, the histidine is adjacent to a disulphide bridge and use was therefore, made of the cysteic acid diagonal technique. To this end, 1 μmole of TLCK inhibited SGT and a control of 1 μmole of native enzyme was digested with pepsin as described in Methods, Chapter II, and subjected to electrophoresis at pH 6.5. Guide strips

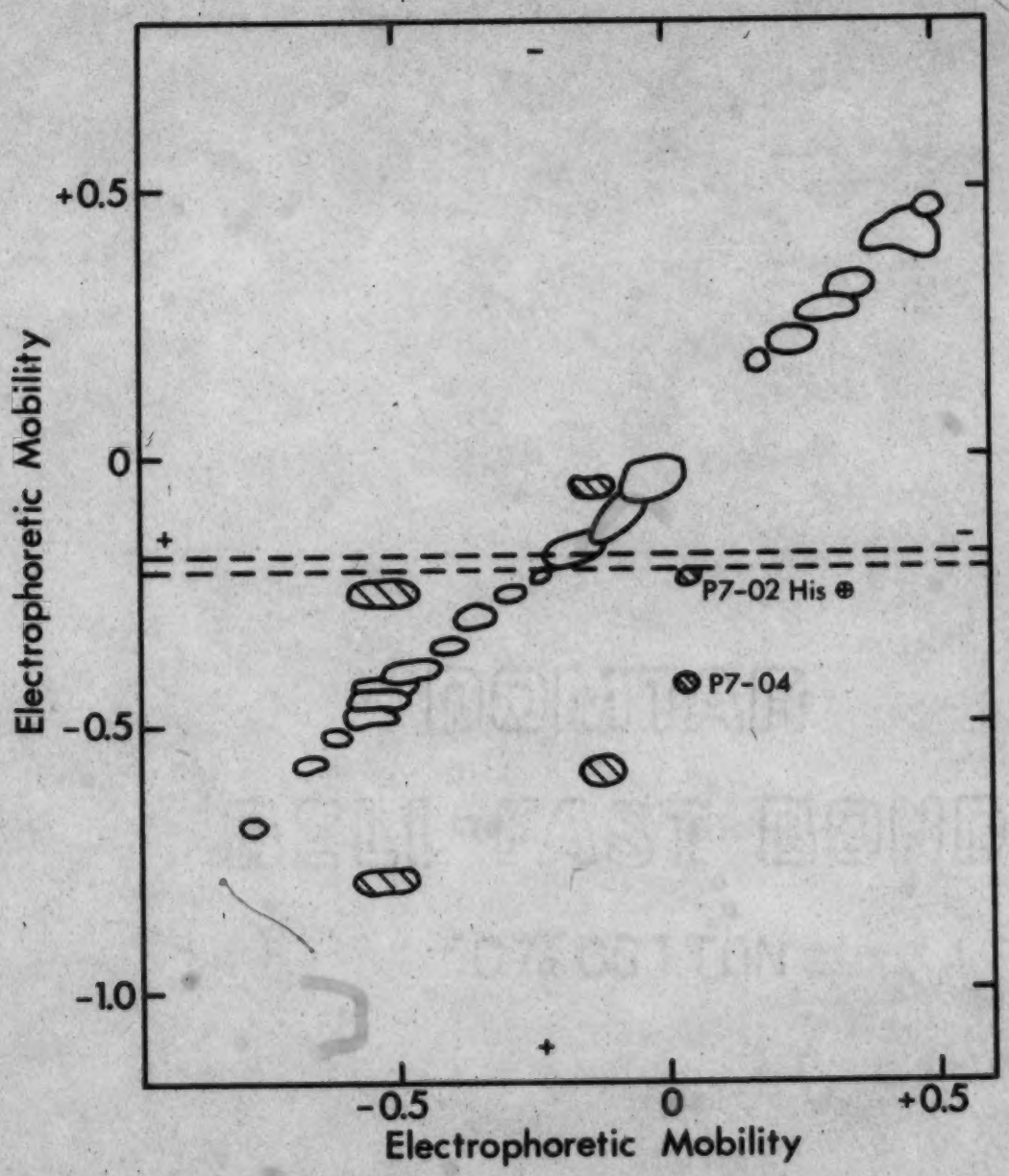


Fig. 29: A pH 6.5/pH 6.5 diagonal peptide map of a peptic digest of SGT. The six cysteic acid peptides moving off the diagonal are indicated by hatching. The Pauly positive P7-02 peptide is indicated by His ⊕.

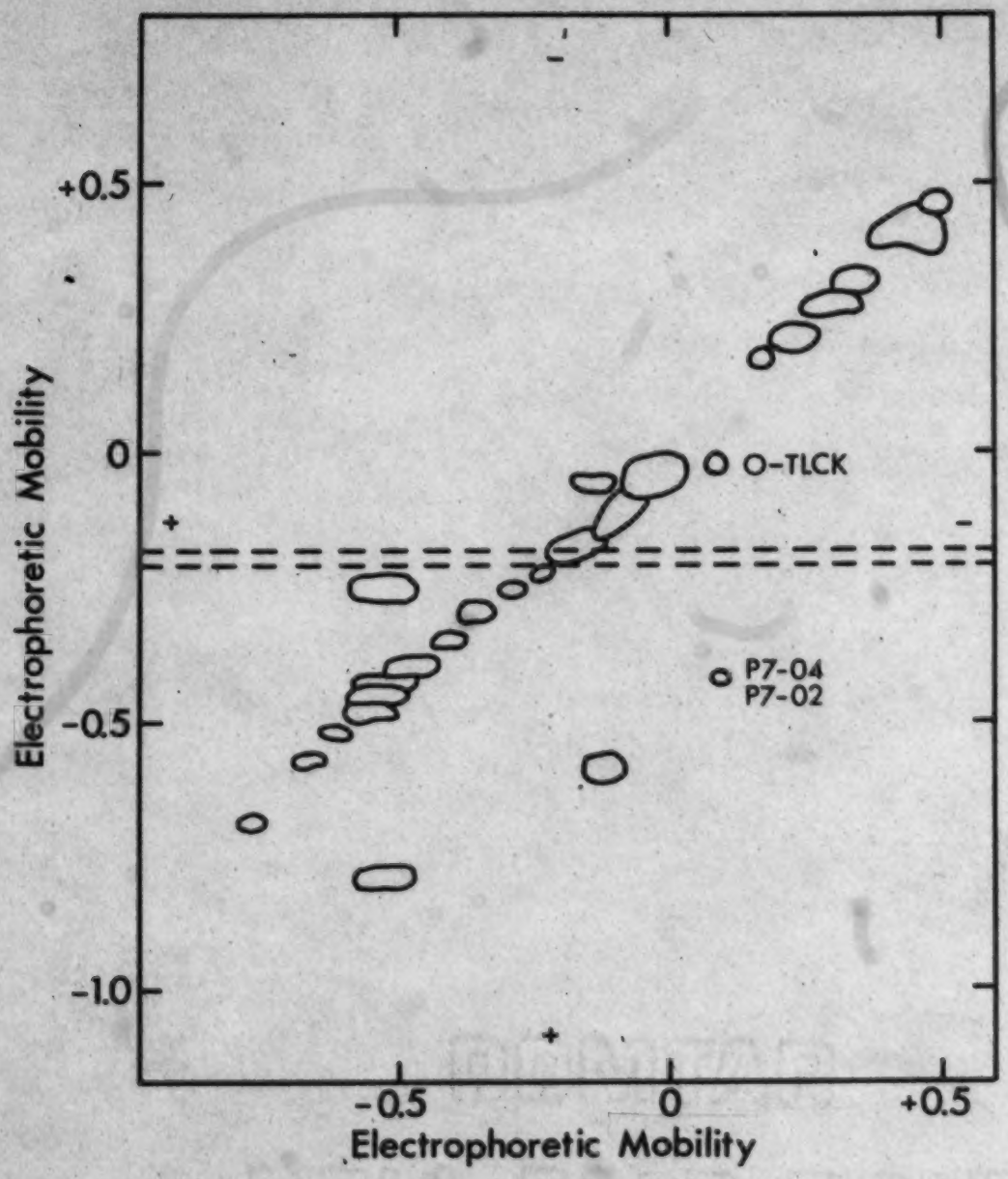


Fig. 28: A pH 6.5/pH 6.5 diagonal peptide map of a peptic digest of TLCK inhibited SGT. O-TLCK and P7-04, P7-02 indicate the positions of ninhydrin positive spots referred to in the text.

were removed from each of these electropherograms and sewn onto new sheets of Whatman 3 MM after oxidation in the fumes of performic acid as described earlier. Each strip was again subjected to electrophoresis under the same conditions but in a direction perpendicular to the original. Individual peptide maps were stained with cadmium-ninhydrin or for histidine with the Pauly procedure (115). The results of this diagonal technique are shown in Figures 28 and 29. Those peptides which moved off the diagonal and approximated the position of the histidine positive peptide in the control electropherogram were noted and the corresponding strip removed from the initial pH 6.5 electropherogram for performic acid oxidation. The peptide designated as O-TLCK was eluted from the subsequent electropherogram, as it most closely approximated the position of the previously observed histidine positive peptide in the control electropherogram. Amino acid analysis indicated that no peptide material was present in the ninhydrin-positive band designated O-TLCK. However, subsequent electrophoresis of band P7-04,P7-02 at pH 1.8, separated this material into two ninhydrin-positive bands P7-04 and P7-02. After subjecting the isolated P7-02 peptide to acid hydrolysis and amino acid analysis, the following composition and NH_2-terminal was found.

(Thr,	Ala,	CM-His,	Cya,	Val,	Ser,	Gly,	Asn)
0.96	2.13	0.98	0.93	0.98	1.89	2.10	2.02

TABLE XV

VARIATION IN HISTIDINE VALUES AFTER INHIBITION WITH TLCK

Uninhibited Enzyme			TLCK Inhibited Enzyme		
Amino Acid	Molar Ratio Expected	Molar Ratio Found	Amino Acid	Molar Ratio Expected	Molar Ratio Found
Lysine	7	7.0	Lysine	7	6.7
Histidine	1	1.4	Histidine	0	0.2
Arginine	9	8.6	Arginine	9	9.1

This is in excellent agreement with the known sequence about the active histidine-57 which is:

Thr-Ala-Ala-His-Cys-Val-Ser-Gly-Ser-Gly-Asn-Asn.

The mate peptide, P7-04, was also isolated and amino acid analyzed, giving excellent agreement with the known sequence and those results previously found in this laboratory (109). The composition results are as follows:

(Ser,	Met (O_2),	Gly,	Cya,	Ala,	Leu).
0.99	0.77	3.06	0.96	1.00	1.00

This is again in agreement with the known sequence about this disulphide bridge residue which is:

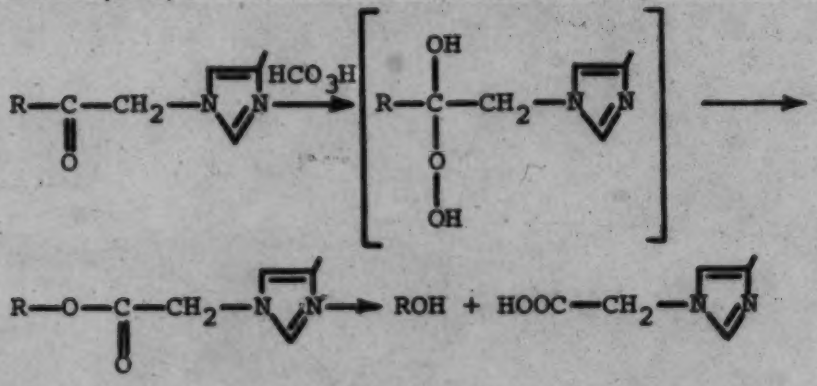
Ser-Met-Gly-Cys-Gly-Gly-Ala-Leu.

When the inhibited enzyme was subjected to amino acid analysis and compared with the native enzyme it was found that the inhibited enzyme apparently lost a histidine as expected from the preceding results. Table XV shows the data for the amino acid analysis of both native and inhibited enzyme.

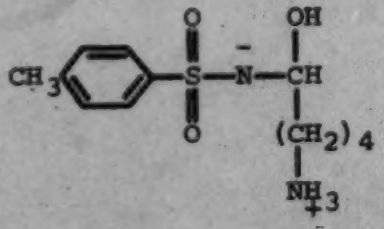
These data, together with those preceding in this section, strongly indicate that, like bovine trypsin, an unmodified histidine-57 is absolutely essential for activity.

As indicated earlier, it was found that the peptide marked O-TLCK, nearest to the position of the histidine positive peptide in the control diagonal at pH 6.5, was found not to be a peptide at all. There is little doubt

that this is a product of oxidative degradation of the modified histidine. Several lines of evidence point to this conclusion. The material, when isolated, was found to be ninhydrin positive and eluted as a single sharp peak in the position of tryptophan on the short column of the amino acid analyzer. Furthermore, according to Stevenson and Smillie (1968), performic acid oxidized alkylated histidine resultant from inhibition of α-chymotrypsin with TPCK, produces 3-carboxymethylhistidine and an alcohol product as a result of a peracid rearrangement (144). The reaction can be diagrammed as shown below.



ROH, in the above reaction, is proposed to be 1-tosylamido-1-hydroxy-5-amino-n-pentane as diagrammed below.



The negative charge on the sulphonamide nitrogen is consistent with the observed mobilities on electrophoresis

of this ninhydrin positive compound at pH 6.5 and pH 1.8. (The mobility at pH 6.5 relative to aspartic acid was -0.1; the mobility at pH 1.8 relative to arginine was +0.9). Indeed, Kenner has reported a sulphonamide nitrogen with a pK_a value as low as 2.3 (159). Although these results do not allow for an absolute conclusion regarding the structure of this by-product, they are all consistent with the hypothesized structure shown previously.

L. Kinetic Isotope Effect of Deuterium Oxide on SGT Hydrolysis of BAEE

The object of this study was to show that SGT, like serine proteases such as α-chymotrypsin (160) and α-lytic protease (76), is affected by deuterium oxide such that the general rate of reaction is decreased two to three fold relative to that in water. According to Bender et al (1962), the rate of a reaction requiring a rate limiting proton transfer is decreased in deuterium oxide, whereas the rate of a reaction subject to nucleophilic catalysis should be affected to only a minor extent (161). In other words, the effect of deuterium oxide is to distinguish between a kinetic mechanism involving a rate limiting proton transfer and one which does not.

The study was carried out in the pH-stat as described in Chapter II, Methods section I-2, and the results of the comparative rate studies over a pH and pH^2

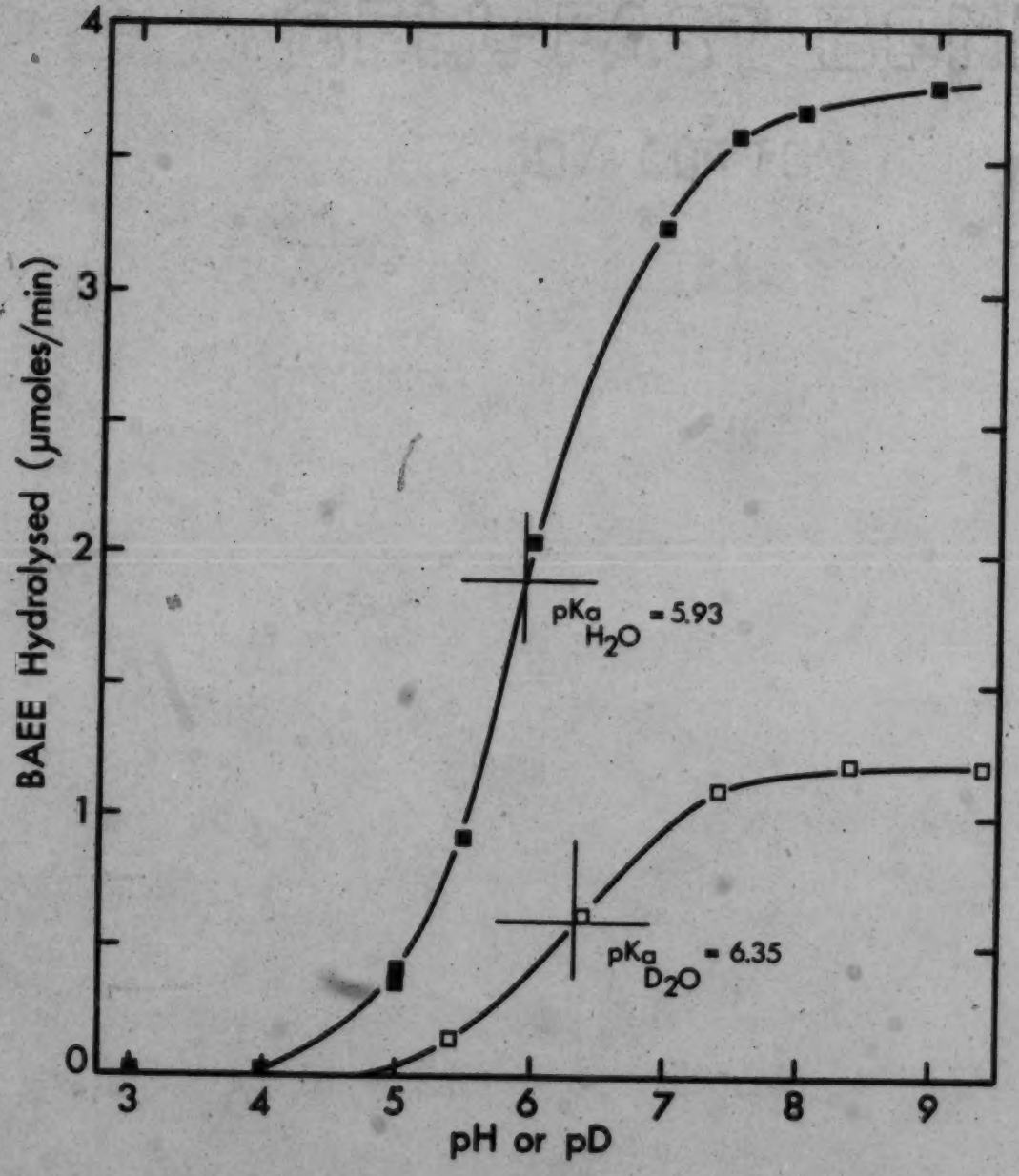


Fig. 30: Plots of the pD and pH dependence of SGT activity as measured with a nitrogen purged pH-stat using BAEE as substrate dissolved in either deuterium oxide or water. □—□ , D_2O; ■—■ , H_2O.

range of 6 units are shown in Figure 30. These data agree with the general concept of a charge transfer complex where a rate limiting proton shift occurs in the active site. However, any dogmatic conclusion regarding mechanism of enzyme action cannot be made from such studies. For example, in D_2O, the pK_a values of ionizing species, the pH, and the solvation of reactants and transition states may differ from those in water. It is important however, that SGT exhibits the same phenomenon in D_2O as do other well characterized serine proteases.

2. STRUCTURAL STUDIES

A. Introduction

In order to provide a more unequivocal proof of the general similarity of Streptomyces griseus trypsin with bovine trypsin and the other serine proteases, the preceding physico-chemical results were augmented with studies directed towards elucidation of the SGT primary structure. As indicated previously, these investigations were initiated by Dr. L. Jurasek of this laboratory in 1968. The contributions of Dr. Jurasek before and after 1970, when this investigator joined him in this endeavor, are outlined in the following introduction.

From peptic digests of the native enzyme, 28 unique amino acid sequences accounting for approximately 95% of the molecule were elucidated by Dr. Jurasek. Among these were the sequences about the active serine and histidine, the three disulphide bridges and the cationic binding site for trypsin substrates. The sequence of residues 1 to 12 at the NH_2-terminal end of the molecule were also elucidated (109,167). These regions showed a remarkable homology with corresponding regions in the amino acid sequence of bovine pancreatic trypsin.

Dr. Jurasek also carried out tryptic digestions of reduced and S-β-aminoethylated SGT. The soluble peptides were recovered and the insoluble residue redigested with

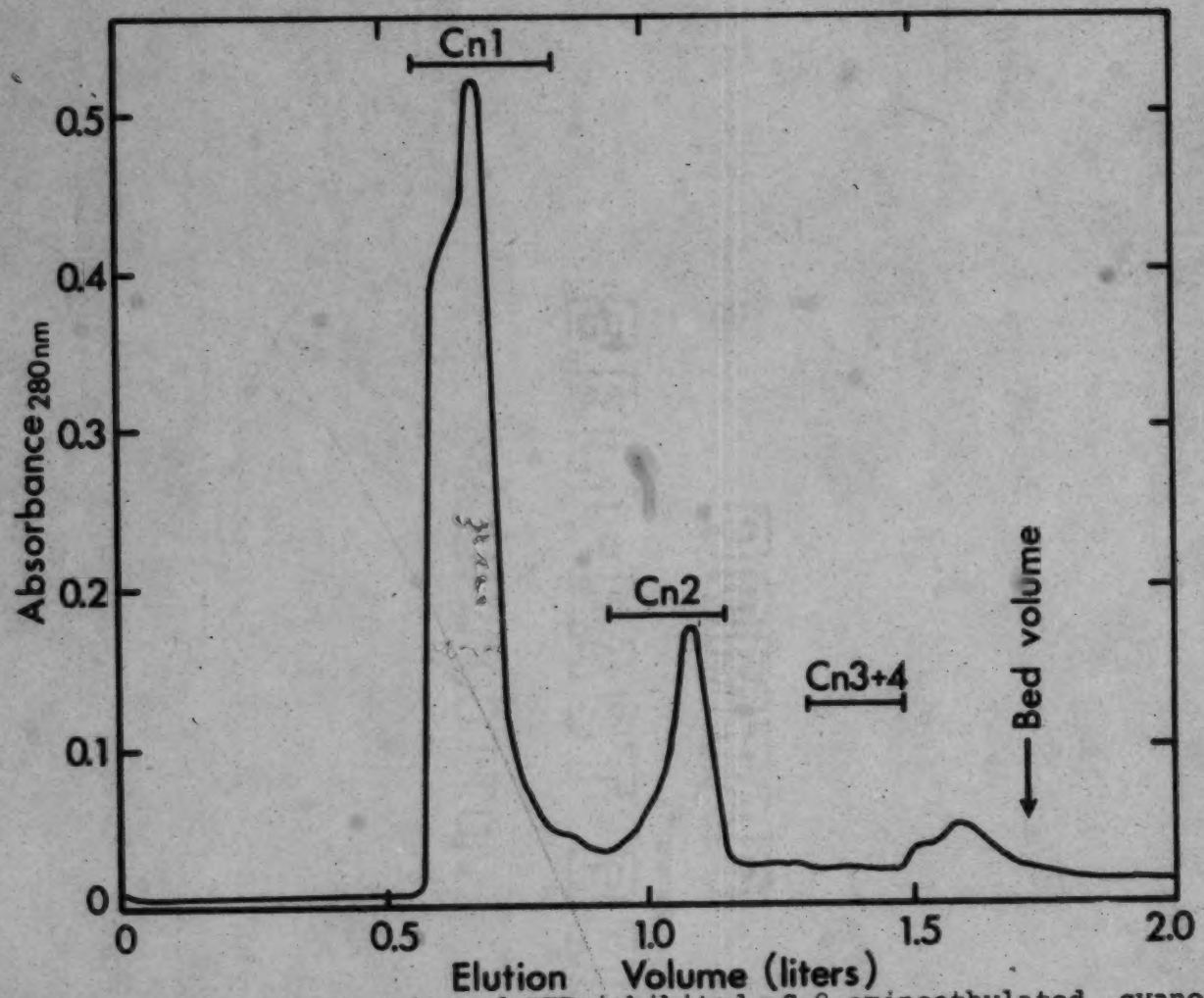


Fig. 31: Sephadex G-50 gel filtration of DFP inhibited, S-β-aminoethylated, cyanogen bromide cleaved SGT in 0.5 M acetic acid. Elution positions of the four fragments are indicated as Cn1, Cn2, and Cn3+4.

chymotrypsin. Following recovery of the soluble fraction, the insoluble portion was in turn redigested with α-lytic protease of Myxobacter 495. The three groups of soluble peptides were separately subjected to ion exchange chromatography on the Technicon peptide analyzer and to final purification by high voltage paper electrophoresis. Characterization of the soluble tryptic peptides provided unique sequences accounting for 65% of the SGT molecule. Peptides obtained from the redigests of the insoluble material accounted for an additional 20%. These sequences provided overlaps for many of the previously sequenced peptic peptides. They also provided the sequences about each of the three methionine residues of the protein. With this information in hand, it should be possible to align fragments arising from cyanogen bromide cleavage of the protein. For this reason and to provide further information on the total sequence, the DFP-treated, reduced and S-β-aminoethylated protein was treated with cyanogen bromide and lyophilized. The freeze-dried fragments were dissolved in 0.5 M acetic acid and applied to a 5 x 88 cm column of Sephadex G-50 fine equilibrated with the same solvent. Fractions were monitored by measuring the optical density at 280 nm and by spotting small portions on paper for high voltage electrophoresis. In this way, three major fractions were observed as indicated in Figure 31. Fraction Cn-3,Cn-4 was found to

TABLE XVI

ESTABLISHED SEQUENCE* OF SGT SHOWING MISSING OR AMBIGUOUS REGIONS

```
                  5                  10                  15                  20
Val-Val-Gly-Gly-Thr-Arg-Ala-Ala-Gln-Gly-Glu-Phe-Pro-Phe-Met-Val-Arg-Leu-Ser-Met-Gly-Cys-

        25                  30                  35                  40
Gly-Gly-Ala-Leu-Tyr-Ala-Gln-Asp-Ile-Val-Leu-Thr-Ala-Ala-His-Cys-Val-Ser-Gly-Ser-Gly-Asn-

45                                              65                  70
Asn- - - - - - - - - - - - - - - -Val-Lys-Val-Arg-Ser-Thr-Lys-Val-Leu-Gln-Ala-Pro-Gly-Tyr-

75                  80                  85                  90                  95
Asn-Gly-Thr-Gly-Lys-Asp-Trp-Ala-Leu-Ile-Lys-Leu-Ala-Gln-Pro-Ile-Asn-Gln-Pro-Thr-Leu-Lys-

                                                                   130
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Lys-Ala-Asn-Val-Pro-Phe-

    135                 140                 145                 150                 155
Val-Ser-Asp-Ala-Ala-Cys-Arg-Ser-Ala-Tyr-Gly-Asn-Glu-Leu-Val-Ala-Asn-Glu-Glu-Ile-Cys-Ala-

                160                 165                 170                 175
Gly-Tyr-Pro-Asp-Thr-Gly-Gly-Val-Asp-Thr-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Met-Phe-Arg-Lys-

        180                 185                 190                 195
Asp-Asn-Ala-Asp-Glu-Trp-Ile-Gln-Val-Gly-Ile-Val-Ser-Trp-Gly-Tyr-Gly-Cys-Ala-Arg-Pro-Gly-

200                 205                 210                 215                 220
Tyr-Pro-Gly-Val-Tyr-Thr-Glu-Val-Ser-Thr-Phe-Ala-Ser-Ala-Ile-Ala-Ser-Ala-Ala-Arg-Thr-Leu
```

*Sequence analysis of SGT as established by Dr. L. Jurasek from peptic and tryptic digests. The disulphide bonds have been found to bridge Cys 22 to Cys 37; Cys 139 to Cys 154 and Cys 166 to Cys 195 (109).

contain varieties of only two unique peptides, and on the basis of amino acid sequence analysis coupled with the previously established amino acid sequences, peptides Cn-4 and Cn-3 were established as being derived from the NH_2-terminal twenty residues of the SGT molecule. Fraction Cn-2, which contained no homoserine or its lactone, was demonstrated to represent the COOH-terminal forty-seven residues of the molecule. This fragment which has an NH_2-terminal phenylalanine residue was sequenced by automatic Edman degradation in the Beckman sequencer and by the characterization of fragments obtained from tryptic, chymotryptic and α-lytic digests. The fraction Cn-1, which eluted at the void volume, was assumed to contain the remaining cyanogen bromide fragment and to be of relatively high molecular weight. Although amino acid and NH_2-terminal analyses were consistent with this assumption, there were indications that this fragment was contaminated with uncleaved protein and possibly aggregates of this fragment with the COOH-terminal fragment Cn-2.

A summary of the sequence data accumulated by Dr. Jurasek is presented in Table XVI. In this representation, the numbering scheme for the designation of amino acid residues in the sequence is based on the final structure as elucidated in this thesis. Thus, as will be described, the total number of residues has now been shown to be 221. From Dr. Jurasek's work the cyanogen

bromide fragments could be arranged in the order Cn-4→ Cn-3→Cn-1→Cn-2, in proceeding from the NH_2-terminal to the COOH-terminal end of the molecule. The sequence was known for residues 1 to 45 at the NH_2-terminal end and from residues 128 to 221 at the COOH-terminal end. In addition, there was an extensive 36 residue sequence established which in Table XVI has been placed from residues 61 to 96. In addition, Dr. Jurasek had isolated a number of smaller peptic and tryptic peptides of known sequence which are not accounted for in the sequences presented in Table XVI.

It is clear that the portion of the sequence for which information was incomplete was present in the large cyanogen bromide fragment Cn-1 (residues 21 to 174 in Table XVI). The purpose of the present investigation was therefore, to isolate this fragment in a homogeneous form and after degradation by appropriate means, to characterize the products for the completion of the structural analysis of the SGT molecule.

B. Purification of Cyanogen Bromide Fragment Cn-1

As indicated above, this fraction as recovered from the Sephadex G-50 system (see Figure 31), appeared to be contaminated with high molecular weight components of the cyanogen bromide reaction mixture, namely uncleaved material and aggregates of the other fragments. For this

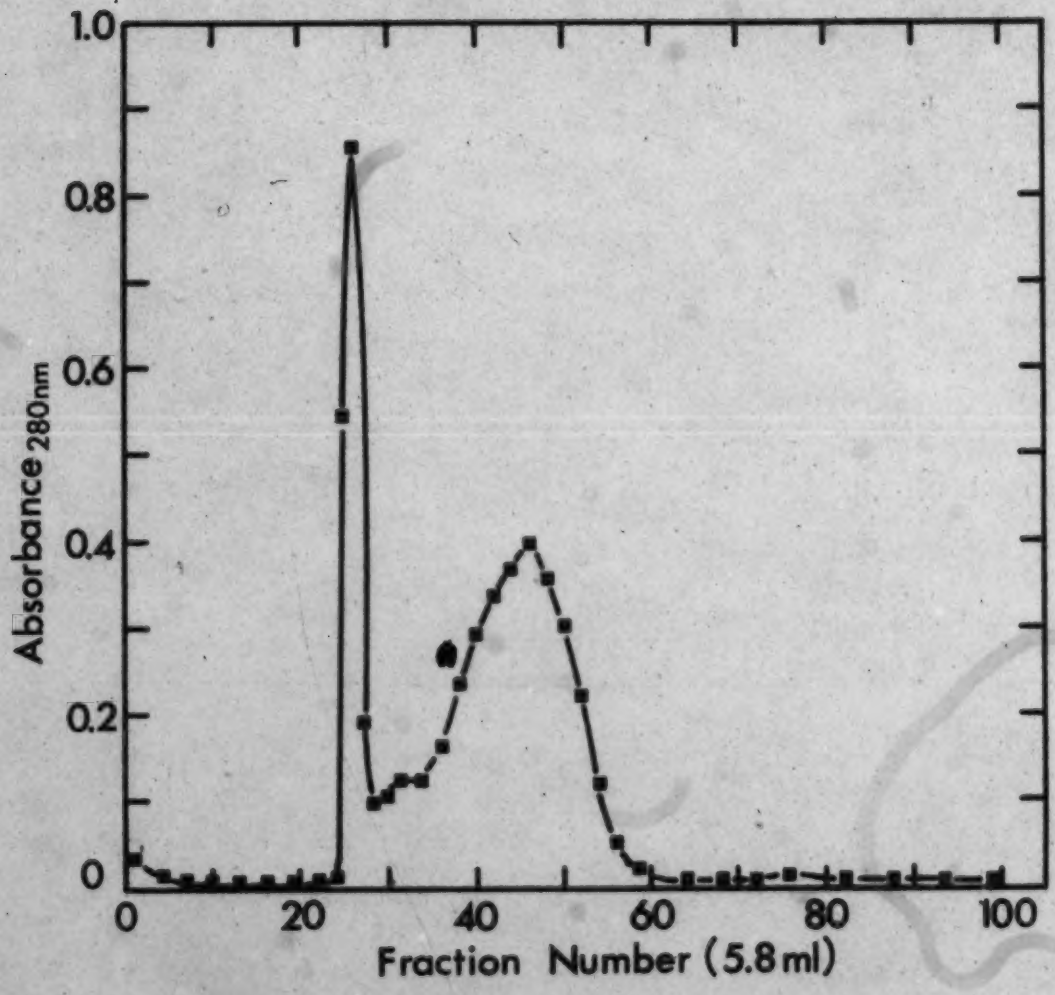


Fig. 32: Sephadex G-75 Superfine gel-filtration of void volume fraction from Sephadex G-50 chromatography shown in Fig. 31. A 100 mg sample was applied to a 2.5 x 100 cm column equilibrated with 50 mM acetic acid and fractions were collected at 20 ml/hr.

reason, the pooled fraction was rechromatographed on Sephadex G-75 Superfine as described in Chapter II, Methods section B1, except that urea was absent and the acetic acid concentration was 50 mM rather than 0.65 M as in later experiments. The results of such an experiment are illustrated in Figure 32. Analysis by the dansyl method, of the first peak eluting at the void volume, gave both DNS-glycine and DNS-phenylalanine. This indicated that it contained both aggregated Cn-1 fragment and Cn-2. A similar analysis of the second broader peak gave only DNS-glycine indicating that it was pure Cn-1 fragment.

Attempts to repeat the isolation of the Cn-1 fragment by the procedures described above met with only limited success both in the hands of Dr. Jurasek and this investigator. In several experiments the fractionation shown in Figure 32 resulted in a single broad peak eluting near the bed volume, where small molecular weight material was expected to elute. These results indicated that proteolytic degradation of the enzyme was occurring during the preparation of the reduced and S-β-aminoethylated protein. This degradation appeared to be attributable to the presence of a contaminating protease in small amounts which was not completely inactivated by treatment with DFP nor by the 8 M urea treatment used in the reduction and S-β-aminoethylation procedure. It was observed that in those experiments in which no measurable proteolytic degradation occurred, the purification of SGT

had involved ion-exchange chromatography on both CM-Sephadex and SE-Sephadex. In those experiments in which extensive autolysis had occurred, the protein had only been purified on CM-Sephadex. The identity of this contaminating protease is unknown but may be Streptomyces griseus protease B which is known from the work of Siegel et al (1972) to retain its activity in the presence of 8 M urea (168). This enzyme elutes close to SGT on both CM-Sephadex and SE-Sephadex (20). Presumably, the purification of SGT on both of these column systems reduces the level of its contamination by this enzyme sufficiently, such that the proteolytic degradation phenomenon is no longer a serious problem. In subsequent work therfore, the SGT used for the preparation of the cyanogen bromide fragments was purified on both the CM- and SE-Sephadex systems.

As indicated previously, a further problem in the purification of the Cn-1 fragment was its apparent aggregation with the Cn-2 fragment and with itself in the acetic acid solvent used in the fractionation of the fragments on Sephadex G-75. This resulted in a poor yield of the purified Cn-1 fragment. In addition, it was observed that the Cn-1 fragment, isolated in 50 mM acetic acid from the Sephadex G-75 column (see Figure 32) was no longer soluble in this solvent after lyophilization. In order to better define the solubility characteristics

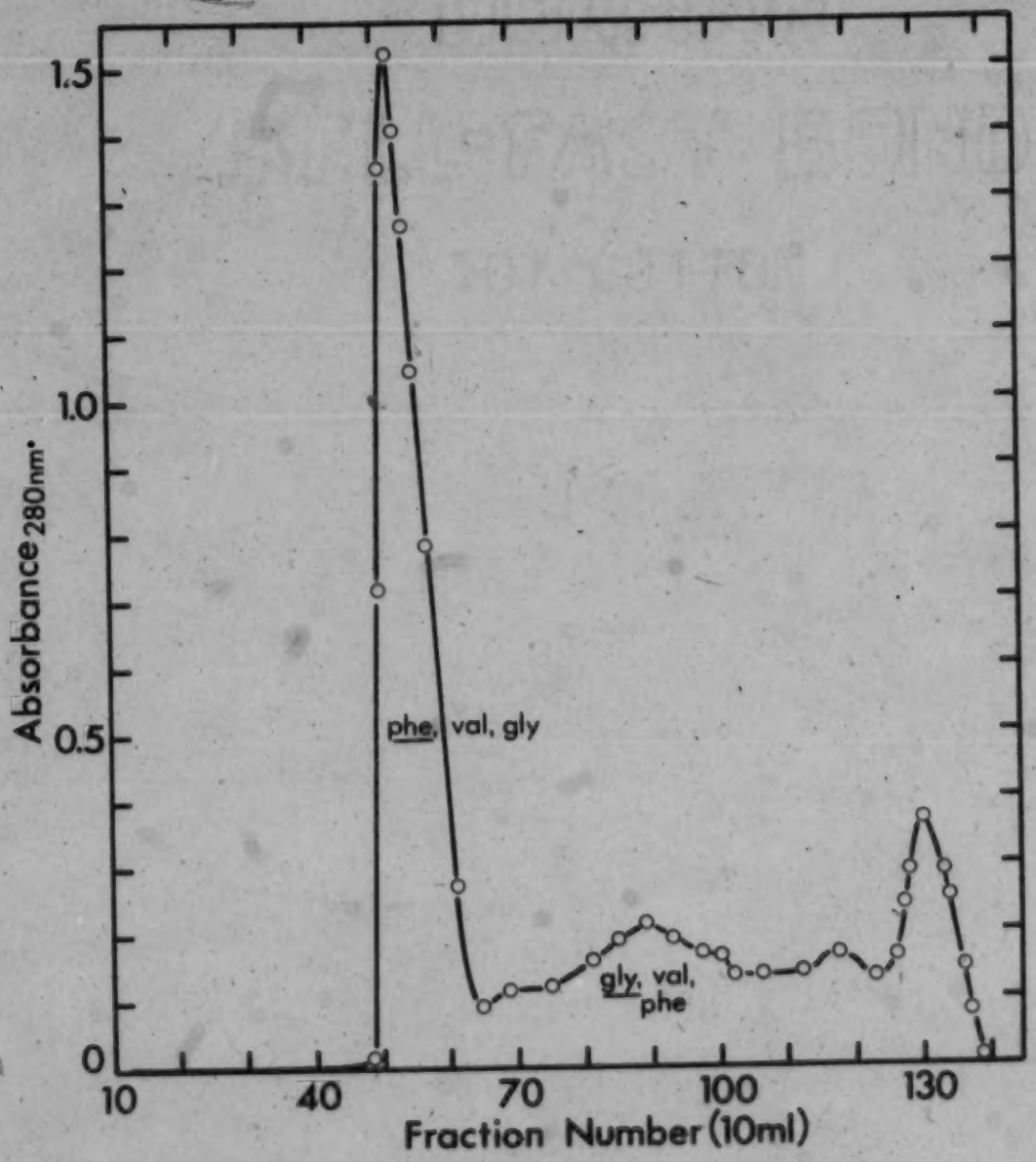


Fig. 33: Sephadex G-75 Superfine gel-filtration of DFP inhibited, S-β-aminoethylated, cyanogen bromide cleaved SGT, previously purified by CM- and SE-Sephadex column chromatography. A 300 mg sample was applied in 55 ml to a 5 x 100 cm column equilibrated with 50 mM acetic acid and pumped at a flow rate of 30 ml/hr. Results of NH_2-terminal determinations on the material in the first two peaks to elute are indicated.

TABLE XVII

QUALITATIVE SOLUBILITY TESTS ON THE CYANOGEN BROMIDE FRAGMENT Cn-1

Solvent	Result
N-Ethylmorpholine, 1 M, pH 8.0	-
N-Ethylmorpholine, 1 M, pH 8.0, 100°	-
Pyridine	-
Aqueous Pyridine	-
Dimethylformamide	-
Dimethylsulphoxide	-
Dioxane	-
Anhydrous Trifluoroacetic Acid	++
Aqueous Trifluoroacetic Acid	-
H_2O / Saturated Urea, Room Temp.	+
Aqueous Urea, 2 M	+
Aqueous Urea, 1 M	+
Aqueous Urea, 0.5 M	+
Glacial acetic acid	-
Glacial acetic / Saturated Urea	++
Acetic acid, 50 mM / Saturated Urea	++
Acetic acid, 50 mM	-

of the fragment for the purpose of devising more satisfactory procedures for its isolation and purification, a series of solubility tests in a number of different solvent systems was carried out. Very small unmeasured fragments of lyophilized peptide were placed in 4 x 25 mm tubes and examined in the designated solvents, using a 16 power monocular lens to observe the apparent degree of dissolution. The results of these qualitative tests are summarized in Table XVII. Although the fragment was insoluble in 50 mM acetic acid, it was found that once dissolved in 50 mM acetic acid—saturated urea at room temperature, the urea could be removed by dialysis against 50 mM acetic acid, leaving the polypeptide in solution. However, the fragment quickly precipitated upon titration of the solution into the alkaline range. These solubility studies formed the basis for the design of subsequent procedures.

Following these tests, several additional attempts were made to prepare the Cn-1 fragment by cyanogen bromide treatment of the S-β-aminoethylated protein. The mixed fragments were routinely dissolved in 50 mM acetic acid—8 M urea and dialyzed against 50 mM acetic acid to facilitate complete solubilization. The mixture was then subjected to fractionation on Sephadex G-75 equilibrated with 50 mM acetic acid. A typical chromatogram is shown in Figure 33 where 300 mg of the mixed fragments

were applied to a 5 x 100 cm column. The major peak which eluted at the position of the void volume contained the majority of the material applied and upon NH_2-terminal analysis gave DNS-phenylalanine, DNS-glycine and DNS-valine. A similar NH_2-terminal result was obtained with material eluted in the second and minor peak (Fractions 73-100) where DNS-glycine was the major product with smaller amounts of DNS-valine and DNS-phenylalanine. These results indicated that two major problems remained. The fact that a majority of the material was eluted at the void volume coupled with the mixed NH_2-terminal results strongly indicated that extensive aggregation among the fragments was occurring in this solvent system (50 mM acetic acid) even though the polypeptide chains remained in solution. Secondly, the finding of NH_2-terminal valine in this material indicated that incomplete cleavage had occurred at methionine-15 and perhaps also at methionine-20 and methionine-174 (see Table XVI). Since valine is the NH_2-terminal residue of both of the small cyanogen bromide fragments Cn-3 and Cn-4, the finding of DNS-valine in this high molecular weight fraction would not be expected unless these two peptides had been only partially cleaved from the NH_2-terminus of Cn-1 by the cyanogen bromide treatment. Incomplete cleavage of the methionines of S-β-aminoethylated proteins by cyanogen bromide treatment has been observed by Schroeder

et al (1967) (168) and presumably results from the conversion, in part, of the methionines to their sulphonium salts by alkylation with ethylenimine.

Because of these unsatisfactory results, alternative procedures for the preparation of the S-β-aminoethylated Cn-1 fragment were explored. Native SGT with disulphide bridges intact was initially treated with cyanogen bromide under the usual conditions of 70% formic acid for 18 hours at room temperature. This procedure should result in complete inactivation of both the SGT activity as well as any contaminating proteases present in the preparation. The reaction should liberate the small fragments Cn-3 and Cn-4 with fragments Cn-1 and Cn-2 linked through the disulphide bridge between residues 166 and 195. In addition, the disulphide bridges in the Cn-1 fragment between residues 139 and 154 should remain intact. Following lyophilization, the Cn-3 and Cn-4 fragments were separated from the high molecular weight disulphide linked Cn-1, Cn-2 fragments by molecular sieving on a Sephadex G-75 Superfine column. The latter fraction was pooled and freeze-dried. Dissolution was accomplished in 120 ml of 8 M urea—0.1 M amediol (2-amino-2-methyl-1,3-propanol) adjusted to pH 3.0 containing 0.8 ml β-mercaptoethanol. However, upon adjustment to pH 9.0 it immediately precipitated. All attempts at its resolubilization in a variety of solvents, including

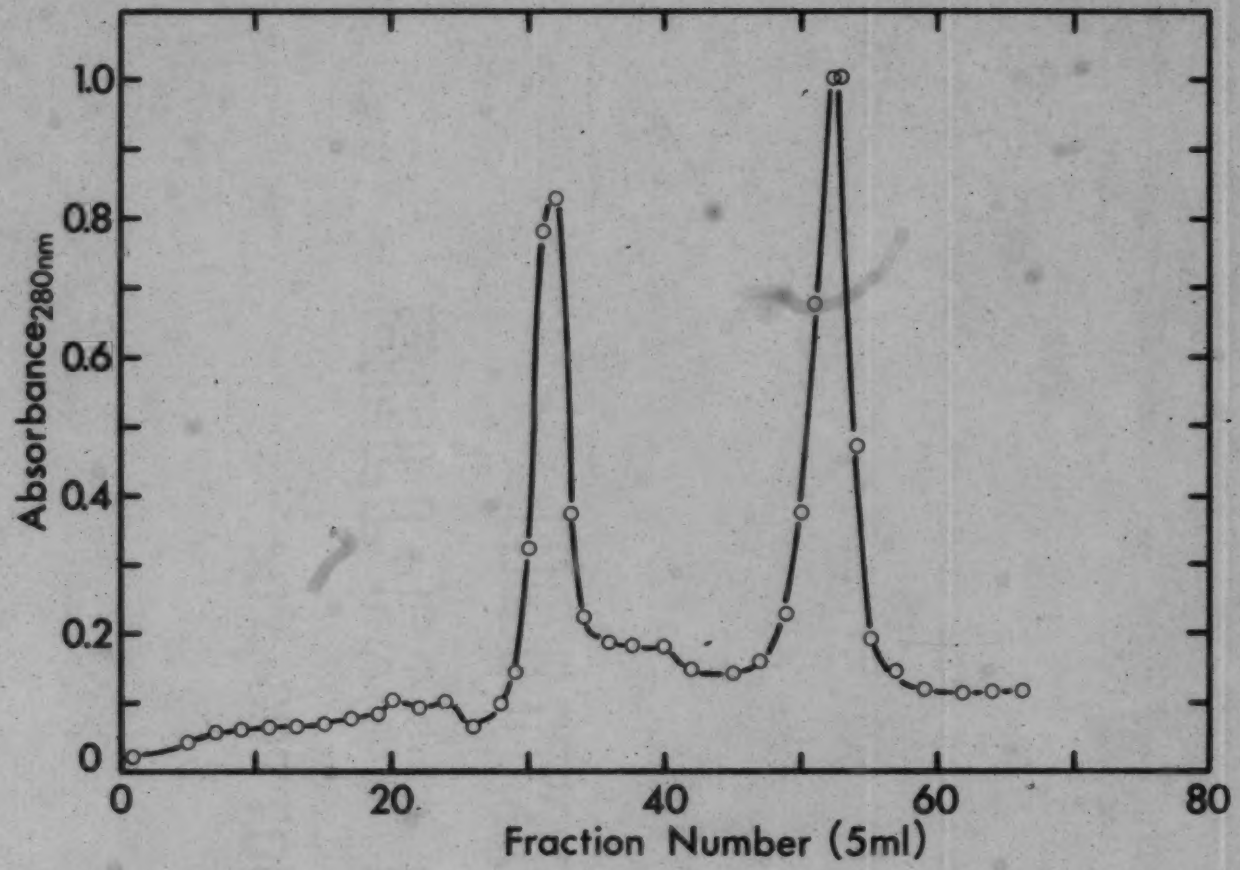


Fig. 34: Sephadex G-75 Superfine gel-filtration of cyanogen bromide cleaved, S-β-aminoethylated SGT. A 25 mg sample in 5 ml was applied to a 2.0 x 100 cm column equilibrated with 0.65 M acetic acid/8 M urea and ran at 8.3 ml/hr.

glacial acetic acid saturated with urea were unsuccessful.

After a number of other unsuccessful attempts to devise an alternative procedure for the preparation of the Cn-1 fragment it was observed that the unfractionated products of the cyanogen bromide cleavage of native SGT were readily soluble in alkaline solutions containing 8 M urea. This solubility behaviour was in sharp contrast to that of the purified disulphide linked Cn-1, Cn-2 fragment. This observation thus permitted the preparation of the reduced and S-β-aminoethylated fragments in a soluble form. The procedure as finally adopted is described in Chapter II, Methods section K2. The mixed S-β-aminoethylated fragments were dissolved in 0.65 M acetic acid—8 M urea at room temperature and applied to a Sephadex G-75 column equilibrated with the same solvent. The increased concentration of acetic acid and the inclusion of 8 M urea in this solvent were designed to avoid the aggregation phenomena observed in previous separations. The results of a trial separation with 25 mg of reduced and S-β-aminoethylated cyanogen bromide fragments using this system are shown in Figure 34. Two major peaks were obtained. NH_2-terminal analysis of the first peak (fractions 29-34) gave DNS-glycine with only traces of impurities, clearly indicating a homogeneous Cn-1 fragment had been obtained. Upon application of the 'dansyl' procedure to the second peak (fractions 46-55)

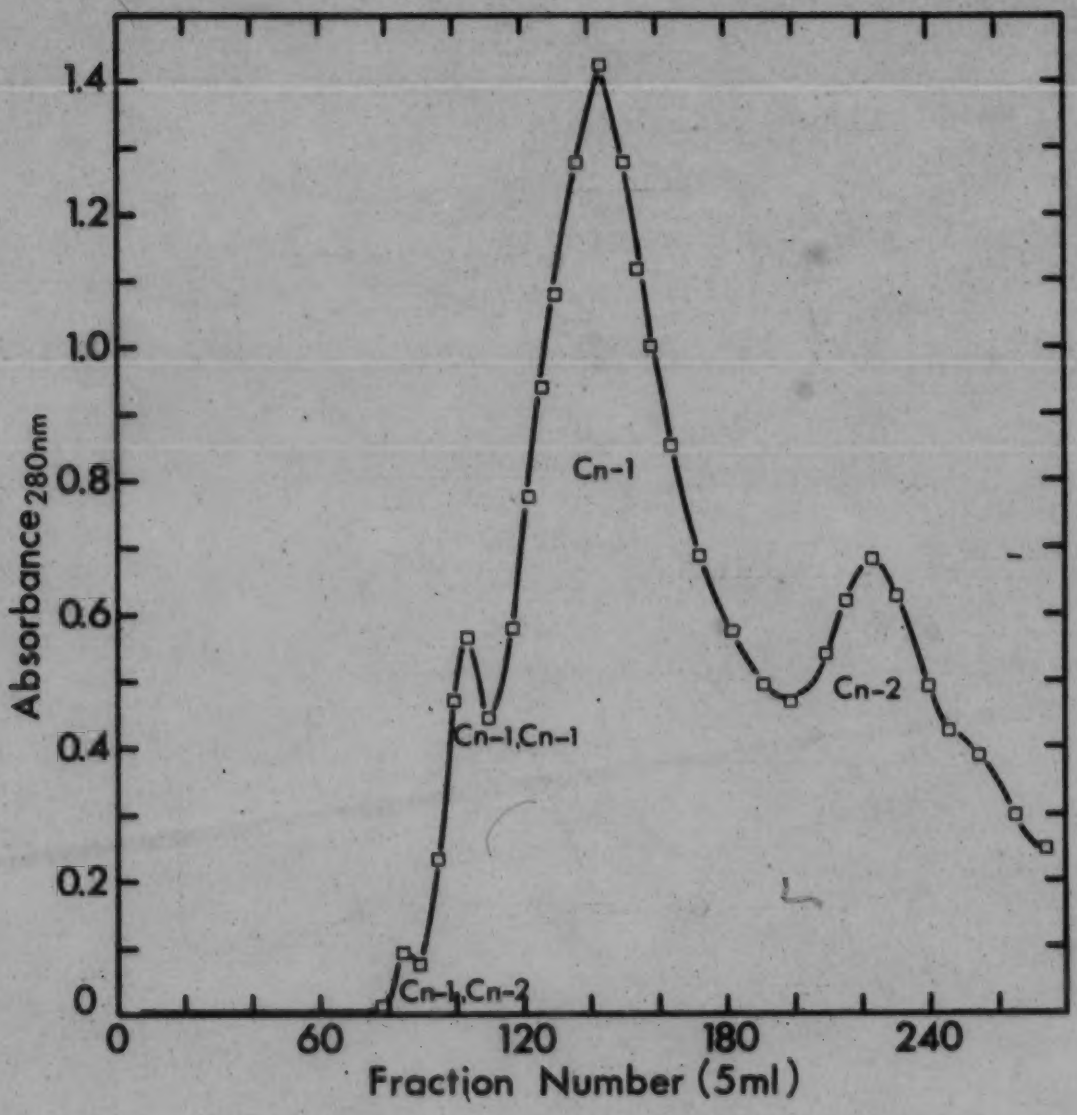


Fig. 35: Sephadex G-75 Superfine gel-filtration of 600 mg of cyanogen bromide cleaved, S-β-aminoethylated SGT on a 5 x 100 cm column equilibrated with 0.65 M acetic acid/ 8 M urea. The column was eluted at 12 ml/hr. Positions of cyanogen bromide fragments and their aggregates, as determined by the 'dansyl procedure', are indicated.

DNS-phenylalanine plus a number of impurities were obtained indicating that the Cn-2 fragment was contaminated, to a degree, with some degradation products and possibly with the Cn-3 and Cn-4 fragments. However, these results indicated that a homogeneous Cn-1 fragment could be prepared by the application of these procedures. The results of scaling up this procedure are illustrated in Figure 35 in which 600 mg of the S-β-aminoethylated cyanogen bromide fragment mixture was applied to a 5 x 100 cm column of Sephadex G-75 Superfine equilibrated with the same solvent. The resolution in this case was less satisfactory and some indications of aggregation between fragments were indicated from the presence of the two small peaks eluting near the void volume. However, NH_2-terminal analysis of pooled fractions (numbers 116-180) gave predominately DNS-glycine with only minor contamination with DNS-phenylalanine and indicated that this preparation of Cn-1 fragment was of satisfactory purity for amino acid sequence analysis. The pooled fractions were dialyzed against 0.65 M acetic acid and lyophilized.

C. Tryptic Digestion of Cyanogen Bromide Fragment Cn-1

Although Dr. Jurasek had been successful in the isolation and sequence analysis of tryptic peptides accounting for a large proportion of the SGT molecule,

portions of the molecule were missing as a result of the insolubility of a number of the larger tryptic peptides in all of the solvent systems examined. It was considered possible that a tryptic digest of the S-β-aminoethylated Cn-1 fragment might result in soluble fragments not previously isolated and thus increase the information concerning the missing regions of the molecule.

Consequently, 3.4 μmoles of Cn-1 peptide were digested under the conditions elucidated in Chapter II, Methods section J2. Unfortunately, the Cn-1 peptide demonstrated poor solubility under the alkaline digestion conditions and formed a finely dispersed suspension which did not appear to decrease in turbidity on reaction. The reaction was halted by adjusting the pH to 4.0 and the product separated into supernatant and insoluble pellet by centrifugation.

The supernatant was concentrated and applied as a 17 cm streak to Whatman 3 MM and electrophoresis at pH 6.5 was performed. After staining guide strips with cadmium-ninhydrin all acidic and basic peptides were further purified by electrophoresis at pH 1.8. The neutral band from the pH 6.5 electrophoresis was also subjected to electrophoresis at pH 1.8. The product of the latter electrophoresis appeared to be four poorly resolved streaking peptides. The neutrals and some poorly resolved acidics were all subjected to electrophoresis at pH 3.5. Of the total digest, sixteen peptides

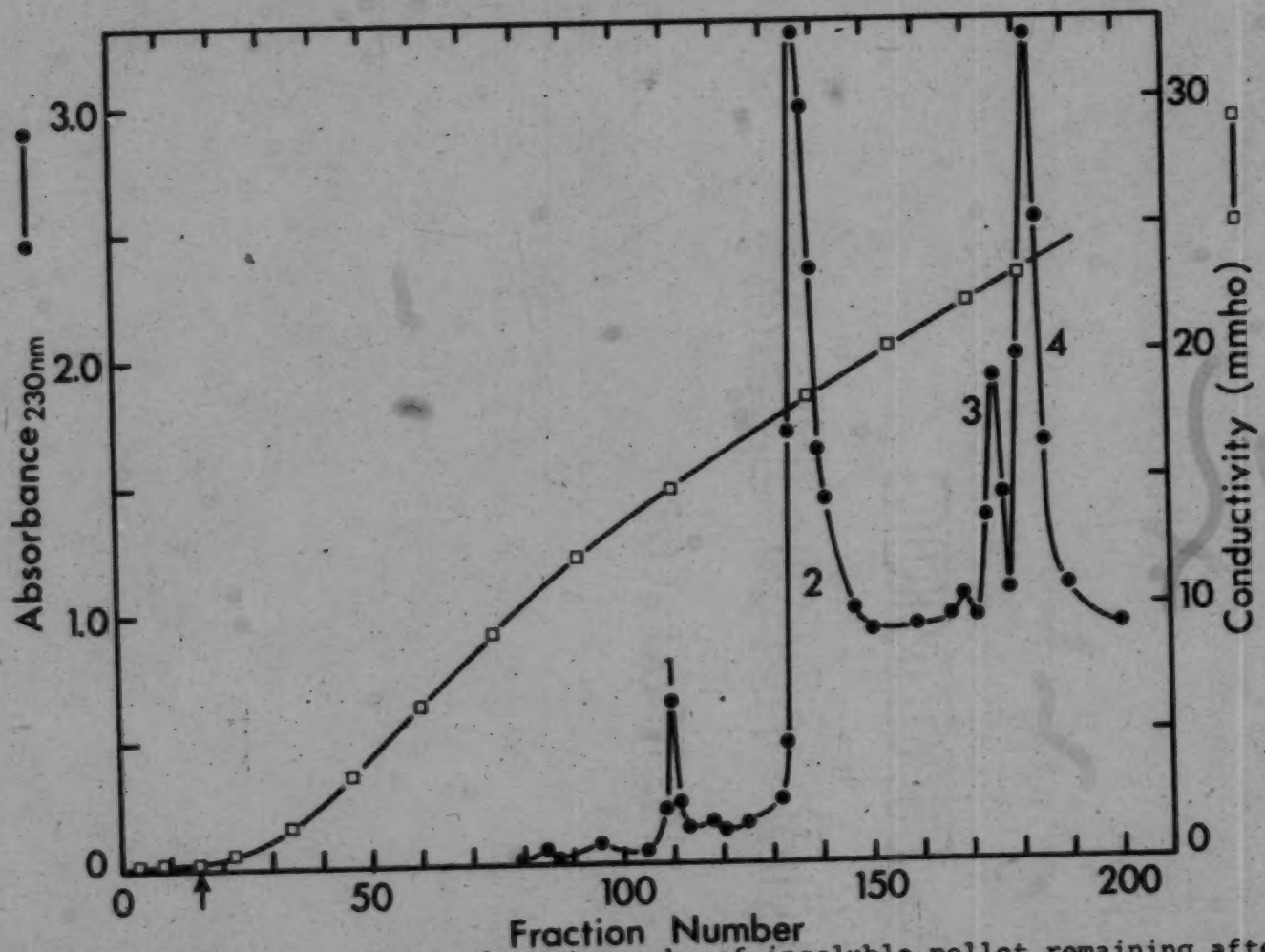


Fig. 36: SE-Sephadex ion exchange chromatography of insoluble pellet remaining after tryptic digestion of Cn-1 fragment. A 2.5 x 100 cm column was equilibrated with 50 mM acetic acid/ 8 M urea, in which the peptides were soluble, and developed with a NaCl gradient (0 to 1.0 M). The arrow indicates the start of the gradient.

appeared pure and were eluted for amino acid analysis and NH_2-terminal determination. In most cases, the peptides were found to be impure by both amino acid analysis and 'dansyl' NH_2-terminal criteria, while those that were identifiable had previously been isolated by Dr. Jurasek. It was obvious, at this time, that the decreased efficiency of digestion incurred by insolubility of the substrate, together with the difficulty of separation of a complex mixture of poorly resolved peptides on paper electrophoresis, resulted in a complete waste of effort.

Nevertheless, it was decided that an attempt to extract some information from the insoluble pellet be made, and to this end it was finally discovered that the pellet could be dissolved in the presence of 0.5 M acetic acid, saturated with urea at room temperature and briefly heated to 60°. The peptides remained in solution on cooling. Chromatography of these peptides on a 2.5 x 100 cm SE-Sephadex C-50 column was performed at room temperature in 50 mM acetic acid/ 8 M urea utilizing a linear NaCl gradient increasing from 0 to 1.0 M. The result is shown in Figure 36. Peaks 1 through 4 were pooled and an attempt was made to desalt the constituent peptides using Sephadex G-10 eluted with 0.5 M acetic acid. In all four fractions the peptides could not be separated from the urea without, what appeared to be, precipitation as demonstrated by tailing

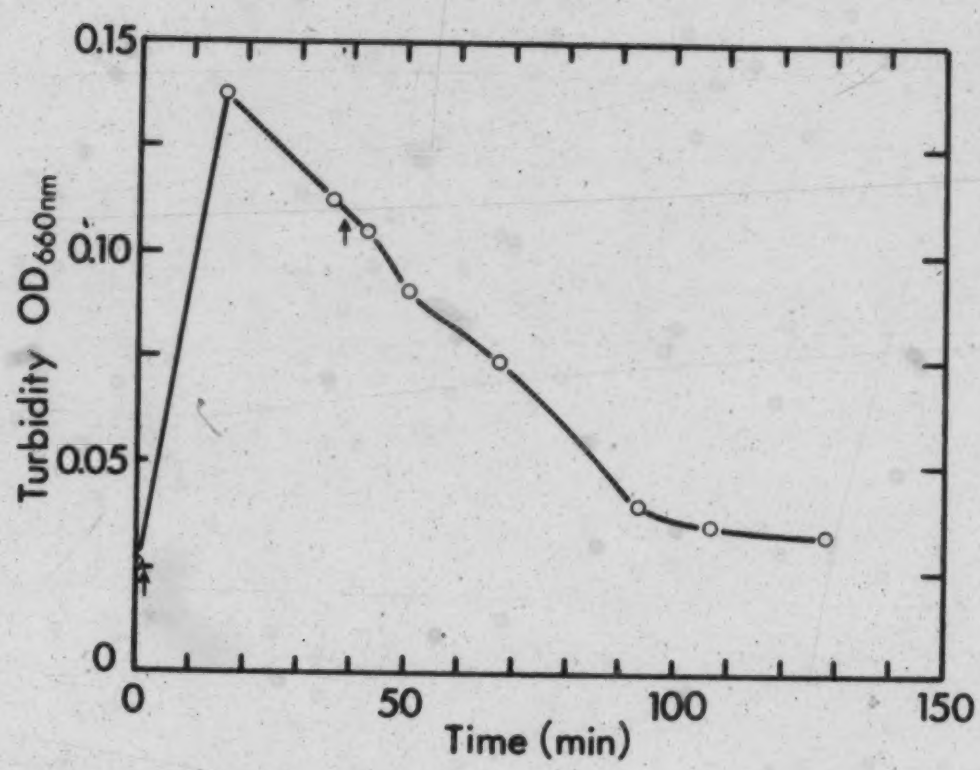


Fig. 37: A plot of variation in turbidity with time on digestion of 300 mg of Cn-1 fragment with α-lytic protease from Myxobacter 495. Times at which enzyme was added are indicated by arrows.

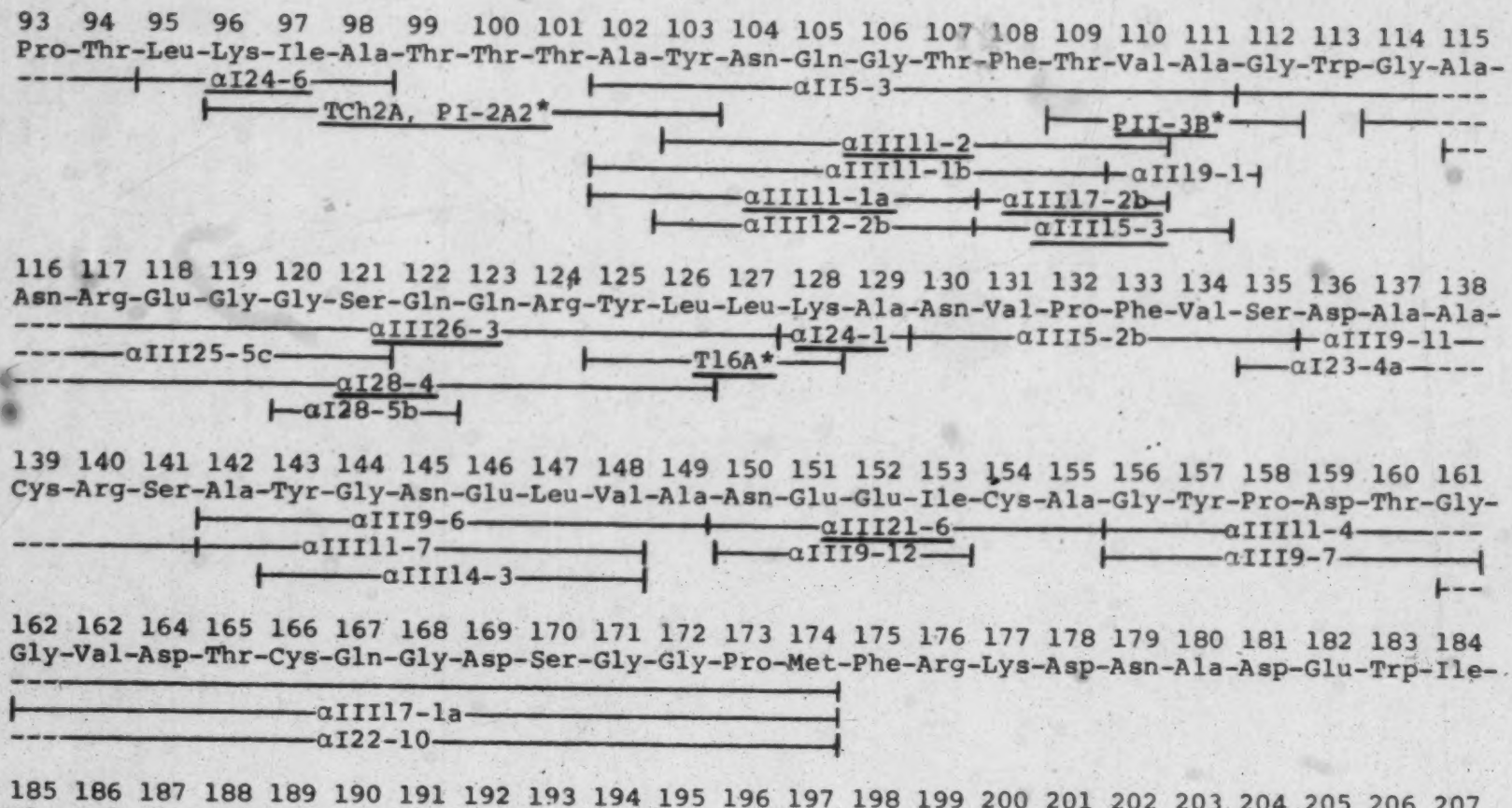
of the fraction throughout the column effluent. Clearly, any method used to desalt these solutions would result in insoluble peptides which could be neither further purified, nor redigested. The information to be gained from tryptic digestion of S-β-aminoethylated Cn-1 or SGT was considered completed and further studies abandoned at this time.

D. α-Lytic Protease Digest of Cn-1 Fragment and Purification of Resultant Peptides

Approximately 300 mg (18.8 μmoles) of S-β-aminoethylated Cn-1 fragment was digested at pH 8.2 as detailed in Chapter II, Methods section J1, and the course of the reaction followed turbidimetrically as illustrated in Figure 37. The protein fragment began to precipitate almost immediately, as indicated by the initial rapid rise in turbidity, but was solubilized as digestion proceeded. The turbidity of the digest mixture at the finish of the reaction was nearly equivalent to that measured at the start and only a negligible precipitate was obtained on centrifugation of the product.

The lyophilized peptides were then dissolved in acetic acid buffer, pH 6.7 and applied to an AG1-X2 (Dowex 1) column. The subsequent stepwise development of this column, as detailed in Chapter II, Methods section

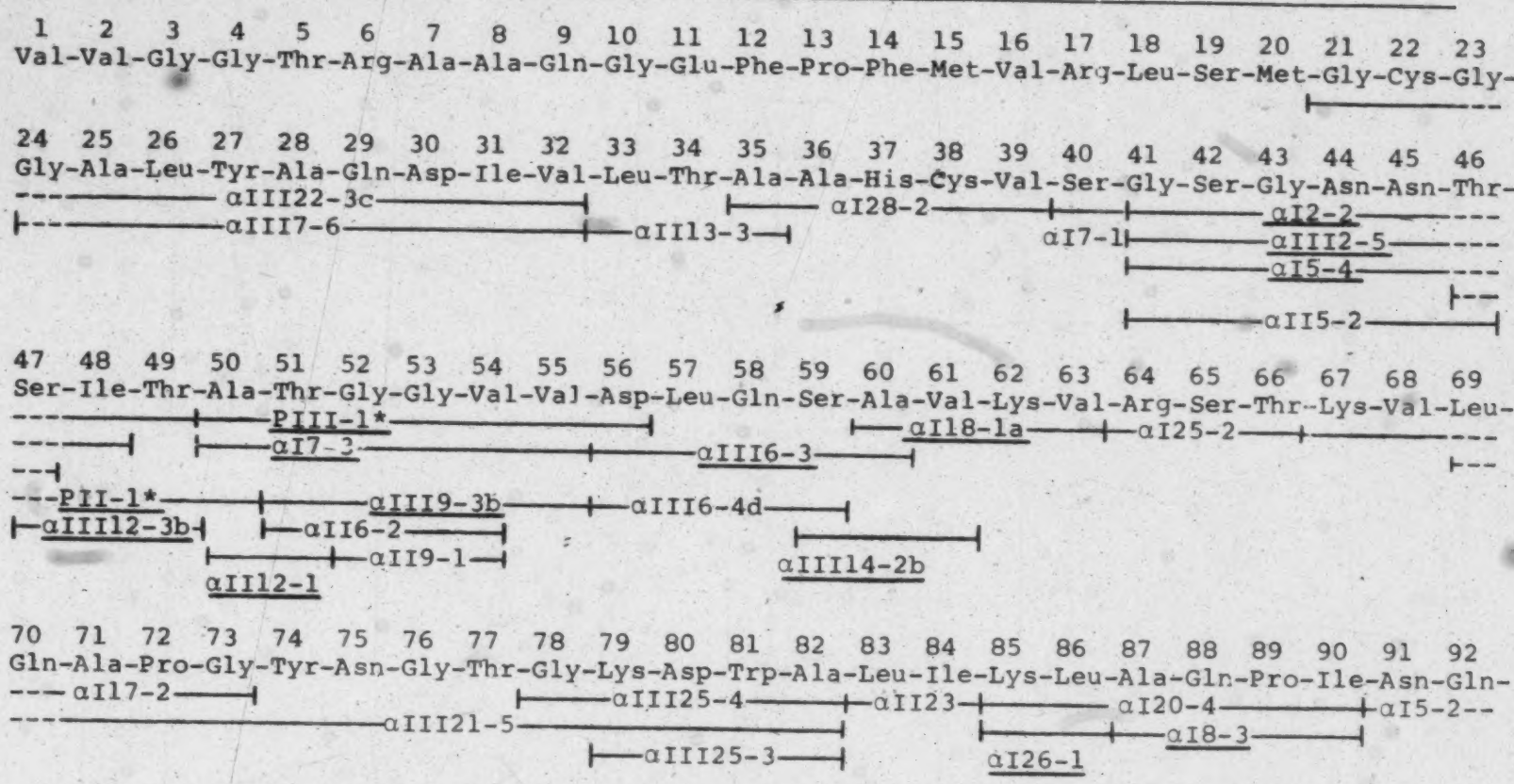
TABLE XIX (cont'd)


93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115
Pro-Thr-Leu-Lys-Ile-Ala-Thr-Thr-Thr-Ala-Tyr-Asn-Gln-Gly-Thr-Phe-Thr-Val-Ala-Gly-Trp-Gly-Ala-
αI24-6
αII5-3
TCh2A, PI-2A2*
PII-3B*
αIII11-2
αIII11-1b
αII19-1
αIII11-1a
αIII17-2b
αIII12-2b
αIII15-3
116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138
Asn-Arg-Glu-Gly-Gly-Ser-Gln-Gln-Arg-Tyr-Leu-Leu-Lys-Ala-Asn-Val-Pro-Phe-Val-Ser-Asp-Ala-Ala-
αIII26-3
αI24-1
αIII5-2b
αIII9-11
αIII25-5c
T16A*
αI23-4a
αI28-4
αI28-5b
139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
Cys-Arg-Ser-Ala-Tyr-Gly-Asn-Glu-Leu-Val-Ala-Asn-Glu-Glu-Ile-Cys-Ala-Gly-Tyr-Pro-Asp-Thr-Gly-
αIII9-6
αIII21-6
αIII11-4
αIII11-7
αIII9-12
αIII9-7
αIII14-3
162 162 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184
Gly-Val-Asp-Thr-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Met-Phe-Arg-Lys-Asp-Asn-Ala-Asp-Glu-Trp-Ile-
αIII17-1a
αI22-10
185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207
Gln-Val-Gly-Ile-Val-Ser-Trp-Gly-Tyr-Gly-Cys-Ala-Arg-Pro-Gly-Tyr-Pro-Gly-Val-Tyr-Thr-Glu-Val-
208 209 210 211 212 213 214 215 216 217 218 219 220 221
Ser-Thr-Phe-Ala-Ser-Ala-Ile-Ala-Ser-Ala-Ala-Arg-Thr-Leu

* Previously Unplaced Peptides of Dr. L. Jurasek

TABLE XIX

SUMMARY OF STREPTOMYCES GRISEUS TRYPSIN SEQUENCE AND ALIGNMENT OF α-LYTIC PEPTIDES FROM CYANOGEN BROMIDE FRAGMENT Cn-1 [a]



[a] Names of sequenced peptides are underlined. eg. αIII14-2b

TABLE XVIII (cont'd)

Peptide	Method of Purification	Mobility pH 6.5	Net Charge	Composition and Sequence (Values are expressed as Mole ratios)	Yield µmoles
αIII21-6	6.5, 1.8	-.43	-1	Asn-Glu-Glu-Ile(AECys,Ala) 1.03 1.07 1.07 0.95 0.83 1.08	0.29
αIII22-3c	6.5, 1.8	0	0	Gly(AECys,Gly,Gly,Ala,Leu,Tyr,Ala,Gln, 0.92 0.61 0.92 0.92 0.97 1.23 0.65 0.97 0.95 Asp,Ile,Val) 1.12 0.62 0.65	0.09
αIII25-3	6.5, 1.8	0	0	Lys(Asp,Trp,Ala) 0.88 1.01 1.11	0.03
αIII25-4	6.5, 1.8	0	0	Gly(Lys,Asp,Trp,Ala) 0.94 0.87 1.02 1.17	0.06
αIII25-5c	6.5, 1.8	0	0	Gly(Ala,Asn,Arg,Glu,Gly,Gly,Ser) 1.07 1.16 1.07 0.83 1.08 1.07 1.07 0.64	0.11
αIII26-3	6.5, 1.8	+.18	+1	Gly-Trp-Gly-Ala-Asn-Arg-Glu(Gly,Gly,Ser, 0.92 ← 0.92 1.19 1.09 0.97 1.06 0.92 0.92 0.89 Gln)Gln-Arg-Tyr-Leu-Leu 1.06 1.06 0.97 1.00 1.01 1.01	0.50

[a] Peptides were fractionated on Dowex 1 and Dowex 50 (peptide analyser) prior to electrophoresis at the pH(s) indicated.

[b] Mobilities were expressed relative to aspartic acid taking its value as -1.0.

[c] According to Offord (1966) (170).

TABLE XVIII (cont'd)

Peptide	Method of Purification	Mobility pH 6.5	Net Charge	Composition and Sequence (Values are expressed as Mole ratios)	Yield µmoles
αIII11-7	6.5, 1.8	-.36	-1	Ala(Tyr,Gly,Asn,Glu,Leu,Val) 1.05 0.91 1.03 1.05 1.06 1.94 0.96	0.90
αIII12-2b	6.5, 1.8	-.06	0	Tyr(Asn,Gln,Gly,Thr) 0.70 1.03 1.03 1.00 0.94	0.10
αIII12-3b	6.5, 1.8	-.09	0	Ser-Ile-Thr 0.92 0.96 1.13	0.26
αIII14-2b	6.5, 1.8	-.06	0	Ser-Ala-Val 0.74 0.99 1.02	0.22
αIII14-3	6.5, 1.8	-.35	-1	Tyr(Gly,Asn,Glu,Leu,Val) 0.79 1.02 1.01 1.02 0.96 0.99	0.32
αIII15-3	6.5, 1.8	-.06	0	Phe-Thr-Val-Ala 0.92 1.02 1.04 1.02	0.32
αIII17-1a	6.5, 1.8	0	0	Gly(Val,Asp,Thr,AECys,Gln,Gly,Asp,Ser, 1.00 0.92 1.03 0.98 0.78 1.01 1.00 1.03 1.08 Gly,Gly,Pro,HSe) 1.00 1.00 0.95 0.66	0.10
αIII17-2b	6.5, 1.8	-.02	0	Phe-Thr-Val 1.03 1.14 0.84	0.01
αIII21-5	6.5, 1.8	-.20	-1	Leu(Glx,Ala,Pro,Gly,Tyr,Asx,Gly,Thr,Gly, 0.95 1.30 1.02 1.17 1.08 0.79 1.08 1.08 0.99 1.08 Lys,Asp,Trp,Ala) 0.81 1.08 1.02	0.02

TABLE XVIII (cont'd)

Peptide	Method of Purification	Mobility pH 6.5	Net Charge	Composition and Sequence (Values are expressed as Mole ratios)	Yield µmoles
αIII9-3b	6.5, 1.8	-.08	0	Thr-Gly-Gly-Val-Val 1.00 0.55 0.55 0.38 0.38	0.24
αIII9-6	6.5, 1.8	-.31	-1	Ala(Tyr,Gly,Asn,Glu,Leu,Val,Ala) 0.96 0.93 1.07 1.05 1.03 0.99 0.99 0.96	0.46
αIII9-7	6.5, 1.8	-.41	-1	Gly(Tyr,Pro,Asp,Thr,Gly) 1.02 0.92 1.03 1.04 0.97 1.02	1.11
αIII9-11	6.5, 1.8	-.67	-1	Asp(Ala,Ala) 1.00 1.00 1.00	2.15
αIII9-12	6.5, 1.8	-.86	-2	(Asn,Glu,Glu,Ile) 0.72 0.87 0.87 0.82	0.12
αIII11-1a	6.5, 1.8	0	0	Ala-Tyr-Asn-Gln-Gly-Thr 0.92 0.80 0.99 0.97 1.13 0.87	0.15
αIII11-1b	6.5, 1.8	0	0	Ala(Tyr,Asn,Gln,Gly,Thr,Phe,Thr) 0.99 0.94 1.04 1.01 1.13 0.95 1.00 0.95	0.34
αIII11-2	6.5, 1.8	-.04	0	Tyr-Asn-Gln-Gly-Thr-Phe-Thr-Val 0.81 1.00 1.03 1.07 0.96 0.98 0.96 0.84	0.05
αIII11-3b	6.5, 1.8	-.08	0	Ser-Ile-Thr-Ala 1.01 0.80 1.05 1.14	0.08
αIII11-4	6.5, 1.8	-.17	-1	Gly(Tyr,Pro,Asp,Thr,Gly,Gly,Val,Asp,Thr, 0.89 0.54 0.87 1.01 1.18 0.89 0.89 1.03 1.01 1.18 AECys,Gln,Gly,Asp,Ser,Gly,Gly,Pro,HSe») 0.47 1.27 0.89 1.01 1.05 0.89 0.89 0.87 0.50	0.03

TABLE XVIII (cont'd)

Peptide	Method of Purification	Mobility pH 6.5	Net Charge	Composition and Sequence (Values are expressed as Mole ratios)	Yield μmoles
αII19-1	1.8	0	0	(Val, Ala) 1.00 1.00	1.36
αII23	1.8	0	0	(Leu, Ile) 0.99 1.02	0.20
αIII2-5	6.5, 1.8	-.37	-1<-2	Gly-Ser-Gly-Asn-Asn*(Thr, Ser, Ile, Thr) 1.01 0.99 1.04 1.04 1.04 0.98 0.99 0.94 0.98	1.05
αIII5-2b	6.5, 1.8	-.08	0	Asn(Val, Pro, Phe, Val, Ser) 1.02 1.04 0.94 1.00 1.04 0.95	0.24
αIII5-3a	6.5, 1.8	-.12	0	(Asn, Gln, Pro, Thr) 1.16 1.14 0.83 0.87	0.73
αIII6-3	6.5, 1.8	-.45	-1	Asp-Leu-Gln-Ser-Ala 1.00 0.72 0.85 0.92 0.87	0.11
αIII6-4d	6.5, 1.8	-.50	-1	Asp(Leu, Gln, Ser) 0.92 1.03 1.09 0.97	0.98
αIII7-2	6.5, 1.8	-.04	0	Ala(Tyr, Asn, Gln, Gly, Thr, Phe, Thr, Val) 1.03 0.78 1.07 1.20 1.18 1.27 0.73 1.27 0.60	0.11
αIII7-6	6.5, 1.8	-.29	-1	Gly(Ala, Leu, Tyr, Ala, Gln, Asp, Ile, Val) 0.92 1.01 1.10 0.95 1.01 1.23 1.26 0.69 0.84	0.15

* Deamidated Asn believed to produce a β-carboxyl peptide bond with adjacent Thr (see text).

TABLE XVIII (cont'd)

Peptide	Method of Purification	Mobility pH 6.5	Net Charge	Composition and Sequence (Values are expressed as Mole ratios)	Yield µmoles
αII9-4	1.8	0	0	Ala(Tyr,Asn,Gln,Gly,Thr,Phe,Thr,Val) 1.04 0.96 0.99 1.04 1.13 0.97 1.02 0.97 0.87	0.56
αII10-2	1.8	0	0	Gly(Gly,Val) 0.96 0.96 1.06	0.11
αII10-4	1.8	0	0	Ala(Tyr,Asn,Gln,Gly,Thr) 1.05 0.74 0.99 1.05 1.18 1.00	0.72
αII11-1	1.8	0	0	Thr 1.00	0.64
αII12-1	1.8	0	0	Ala-Thr 1.00 0.95	0.71
αII13-3	1.8	0	0	Leu(Thr,Ala) 0.81 1.10 0.98	0.25
αII17-1	1.8	0	0	(Ile,Thr) 1.00 0.95	1.36
αII18-1	1.8	0	0	Ala 1.00	0.67
αII18-2	1.8	0	0	(Ser,Ala) 0.92 1.00	0.09
αII18-4	1.8	0	0	(Ser,Ile,Thr) 0.75 0.96 1.04	0.03

TABLE XVIII (cont'd)

Peptide	Method of Purification	Mobility pH 6.5	Net Charge	Composition and Sequence (Values are expressed as Mole ratios)	Yield μmoles
αII3-2	1.8	0	0	Gly(Ser,Gly,Asn,Asn,Thr,Ser) 0.99 0.87 0.99 1.01 1.01 0.95 0.87	0.97
αII4-1	1.8	0	0	Ala(Thr,Gly,Gly,Val,Val) 0.98 0.95 1.03 1.03 0.59 0.59	1.64
αII5-1	1.8	0	0	Ala(Gln,Pro,Ile) 1.05 1.00 0.73 0.95	1.63
αII5-2	1.8	0	0	Gly(Ser,Gly,Asn,Asn,Thr) 1.04 1.02 1.04 0.95 0.95 1.00	0.19
αII5-3	1.8	0	0	Ala(Tyr,Asn,Gln,Gly,Thr,Phe,Thr,Val,Ala) 0.89 1.09 0.94 1.11 1.37 0.90 1.13 0.90 0.78 0.89	0.48
αII6-2	1.8	0	0	(Thr,Gly,Gly,Val) 1.07 0.99 0.99 0.94	0.05
αII6-3	1.8	0	0	(Thr,Gly,Gly,Val,Val) 0.97 1.01 1.01 0.54 0.54	0.22
αII9-1	1.8	0	0	Gly(Gly,Val) 0.98 0.98 1.05	0.11
αII9-2	1.8	0	0	Val 1.00	0.07
αII9-3	1.8	0	0	Ala(Tyr,Asn,Gln,Gly,Thr,Ala) 0.89 0.69 1.19 0.86 1.13 0.82 0.89	0.13

TABLE XVIII (cont'd)

Peptide	Method of Purification	Mobility pH 6.5	Net Charge	Composition and Sequence (Values are expressed as Mole ratios)	Yield µmoles
αI24-1	6.5, 1.8	+.64	+1	Lys-Ala 1.02 0.98	1.28
αI24-6	6.5, 1.8	+.41	+1	Leu-Lys-Ile-Ala 1.15 0.94 1.00 0.90	0.54
αI25-2	6.5, 1.8	+.45	+1	Arg(Ser,Thr) 1.12 0.96 0.92	1.06
αI26-1	6.5, 1.8	+.57	+1	Lys(Leu) 0.89 1.11	0.49
αI28-2	6.5, 1.8	+.57	+1<+2	Ala(Ala,His,AECys,Val) 0.97 0.97 1.00 0.64 1.08	0.03
αI28-4	6.5, 1.8	+.21	+1	Ala-Asn-Arg-Glu-Gly-Gly-Ser-Gln-Gln- 1.14 1.04 0.99 1.03 1.15 1.15 0.92 1.03 1.03 Arg(Tyr)Leu 0.99 0.87 0.63	0.18
αI28-5b	6.5, 1.8	0	0	Gly(Ser,Gln) 1.23 0.98 0.80	0.01
αII2-1	1.8	0	0	Asn(Gln,Pro,Thr) 0.99 1.04 0.99 0.99	1.30
αII2-3	1.8	0	0	Asn(Val,Pro,Phe,Val,Ser) 1.21 0.95 0.80 1.00 0.95 1.08	0.60

TABLE XVIII (cont'd)

Peptide	Method of Purification	Mobility pH 6.5	Net Charge	Composition and Sequence (Values are expressed as Mole ratios)	Yield µmoles
αI12-2	6.5, 1.8	-.05	0	Ser-Ala 0.90 1.10	0.06
αI17-2	6.5, 1.8	+.30	+1	Lys(Val,Leu,Gln,Ala,Pro,Gly) 0.85 0.89 1.10 0.95 1.15 0.73 1.07	0.18
αI18-1a	6.5, 1.8	+.40	+1	Ala-Val-Lys-Val 0.98 0.98 1.08 0.98	0.26
αI18-4	6.5, 1.8	-.14	0>-1	Gly(Val,Asp,Thr,AECys,Gln,Gly,Asp,Ser, 0.99 0.98 1.02 0.95 Trace 1.07 0.99 1.02 1.00 Gly,Gly,Pro,HSe) 0.99 0.99 0.93 0.70	0.09
αI20-4	6.5, 1.8	+.31	+1	Lys(Leu,Ala,Gln,Pro,Ile) 0.91 1.19 1.02 0.99 0.82 1.08	0.18
αI22-5	6.5, 1.8	+.41	+1	Leu-Lys-Ile-Ala 1.05 1.11 0.99 0.85	0.08
αI22-10	6.5, 1.8	+.24	+1	Gly(Gly,Val,Asp,Thr,AECys,Gln,Gly,Asp, 0.92 0.92 1.03 1.05 1.03 Trace 1.12 0.92 1.05 Ser,Gly,Gly,Pro,HSe>) 1.14 0.92 0.92 0.83 0.39	0.03
αI23-4a	6.5, 1.8	+.34	+1	Ser(Asp,Ala,Ala,AECys,Arg,Ser) 0.48 1.02 1.16 1.16 0.71 0.99 0.48	0.11

TABLE XVIII

AMINO ACID COMPOSITION AND SEQUENCE OF α-LYTIC PEPTIDES FROM CYANOGEN BROMIDE FRAGMENT Cn-1

Peptide	Method of Purification[a]	Mobility pH 6.5[b]	Net Charge[c]	Composition and Sequence (Values are expressed as Mole ratios)	Yield μmoles
αI2-2	3.5	0	0	Gly-Ser-Gly-Asn-Asn-Thr-Ser(Ile,Thr) 1.04 0.95 1.04 1.06 1.06 0.95 0.95 1.00 0.95	0.28
αI5-2	1.8, 3.5	0	0	(Asn,Gln,Pro,Thr) 0.90 1.14 1.00 0.96	0.11
αI5-4	1.8, 3.5	0	0	Gly-Ser-Gly-Asn-Asn-Thr-Ser 1.00 0.94 1.00 1.09 1.09 0.92 0.94	0.21
αI6-5	1.8	0	0	Asn(Val,Pro,Phe,Val,Ser) 1.14 0.94 0.93 0.93 0.94 1.11	0.03
αI7-1	1.8	0	0	Ser 1.00	0.09
αI7-3	1.8, 3.5	0	0	Ala-Thr-Gly-Gly-Val-Val 1.07 1.02 1.10 1.10 0.85 0.85	0.31
αI8-3	1.8	0	0	Ala-Gln-Pro-Ile 0.96 0.97 0.86 1.07	0.66
αI10-1a	6.5, 1.8	0	0	Ile 1.00	0.03

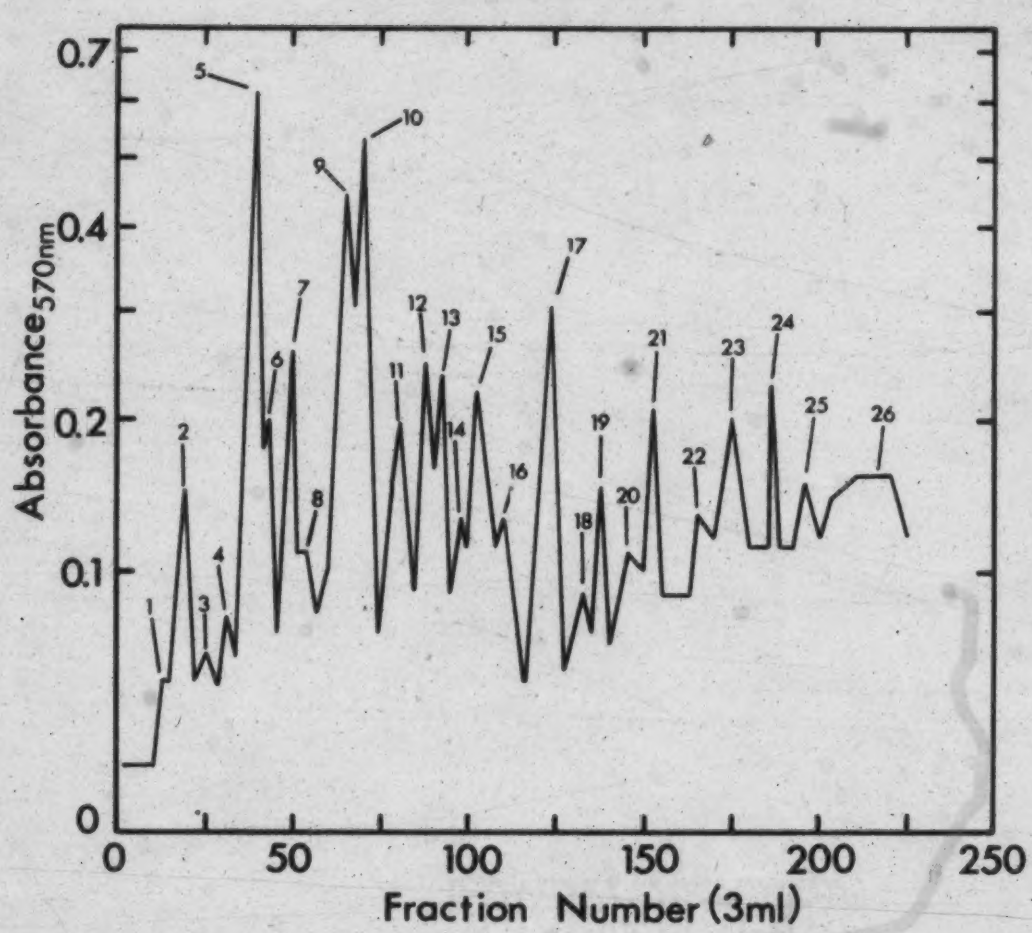


Fig. 40: Chromatography on Chromobead type P resin of fraction III from Dowex 1 fractionation of the α-lytic digest of Cn-1. The flow rate was 25 ml/hr and 3.0 ml fractions were collected. The effluent was monitored by automatic ninhydrin analysis using 7% of the elution volume.

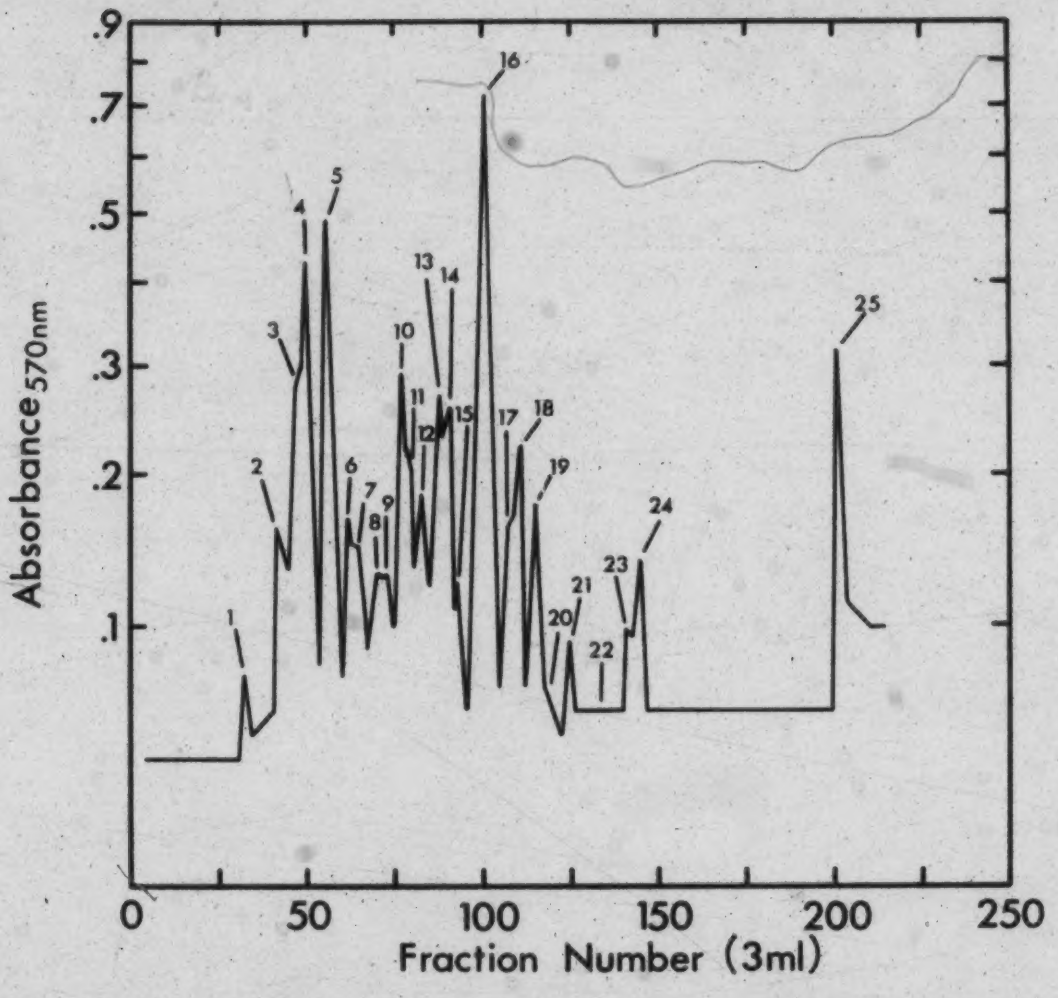


Fig. 39: Chromotography on Chromobead type P resin of fraction II from Dowex 1 fractionation of the α-lytic digest of Cn-1. The flow rate was 25 ml/hr and 3.0 ml fractions were collected. The effluent was monitored by automatic ninhydrin analysis using 7% of the elution volume.

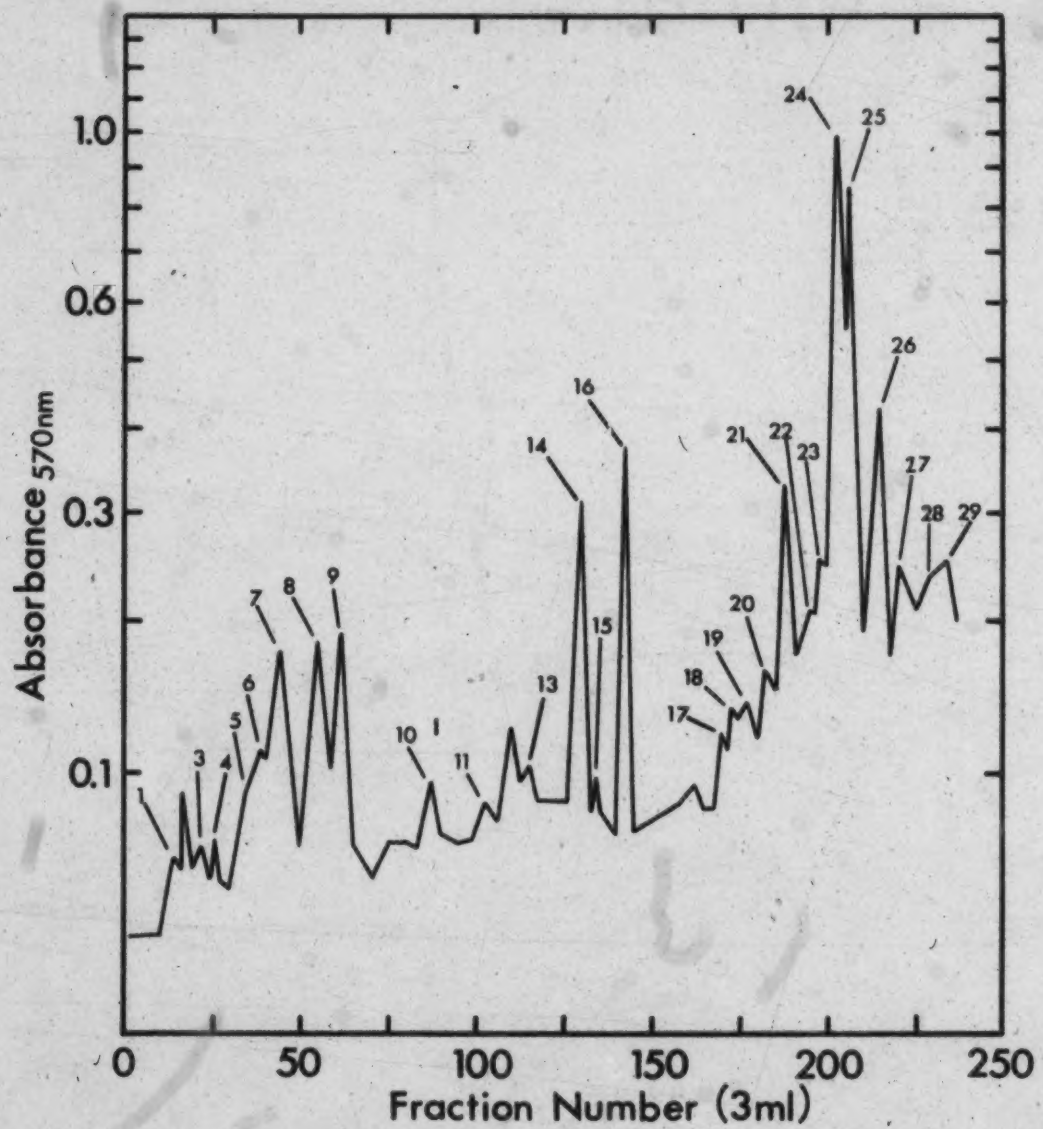


Fig. 38: Chromatography on Chromobead type P resin of fraction I from Dowex 1 fractionation of the α-lytic digest of Cn-1. The flow rate was 25 ml/hr and 3.0 ml fractions were collected. The effluent was monitored by automatic ninhydrin analysis using 7% of the elution volume.

B3, resulted in a partial separation of the digest mixture into basic, neutral and acidic peptides (55). It was found that the basic and acidic fractions contained appreciable amounts of neutral peptides, but the system adequately served the purpose of simplifying subsequent separation procedures. The three fractions were lyophilized immediately.

The next step in the preparative procedure involved fractionation of each of the Dowex 1 fractions on a Technicon autoanalyzer system for peptide chromatography. The system was used as detailed in Chapter II, Methods section B4, and profiles of the chromatography of the basic, neutral and acidic fractions are shown in Figures 38 to 40.

The contents of the peak tubes from the preceding fractionations were pooled, lyophilized and finally dissolved in 2 ml of deionized distilled water. The peptides were then further purified electrophoretically as detailed in Table XVIII. Details, such as electrophoretic mobility at pH 6.5, calculated net charge, molar ratios of residues from amino acid analysis and yields in μmoles are also indicated on the aforementioned table. The sequence already established by Dr. Jurasek was shown previously in Table XVI, while Table XIX indicates the positions of the isolated α-lytic peptides on the known sequence established by Dr. Jurasek and, in

addition, indicates where sequence information has been extended.

It should be mentioned at this time, that the peptide nomenclature in the preceding tables, follows directly from the isolation procedure. Thus αI6-2a indicates a peptide arising from an α-lytic protease digest, where the Roman numeral 'I' indicates that it eluted in the first or basic fraction from the Dowex 1 column procedure; the Arabic numeral '6' indicates the position of elution off the peptide analyzer; the Arabic numeral '2' indicates that this was the second most basic peptide after the first electrophoresis at the designated pH on Table XVIII, while the letter 'a' indicates that it was the most basic of those peptides isolated after performing a second electrophoresis, again at the designated pH on the previous table. In short, 'α' refers to the type of digest, 'I' to the order of elution off the Dowex 1 column, '6' to the order of elution off the peptide analyzer and '2a' to the relative basicity on the first and second electrophoresis, respectively. Accordingly, the neutral fraction eluted from Dowex 1 would use the Roman numeral 'II' and the acidic fraction the Roman numeral 'III'.

The result of the final alignment using unplaced α-lytic digest peptides, plus several unplaced peptides from Dr. Jurasek's work (designated with an asterisk in

TABLE XX

ASSUMED PRIMARY SEQUENCE OF STREPTOMYCES GRISEUS TRYPSIN

```
 1                   5                  10                  15                  20
Val-Val-Gly-Gly-Thr-Arg-Ala-Ala-Gln-Gly-Glu-Phe-Pro-Phe-Met-Val-Arg-Leu-Ser-Met-

                   25                  30                  35                  40
Gly-Cys-Gly-Gly-Ala-Leu-Tyr-Ala-Gln-Asp-Ile-Val-Leu-Thr-Ala-Ala-His-Cys-Val-Ser-

                   45                  50                  55                  60
Gly-Ser-Gly-Asn-Asn-Thr-Ser-Ile-Thr-Ala-Thr-Gly-Gly-Val-Val-Asp-Leu-Gln-Ser-Ala-
                                       ↑1                                  ↑2
                   65                  70                  75                  80
Val-Lys-Val-Arg-Ser-Thr-Lys-Val-Leu-Gln-Ala-Pro-Gly-Tyr-Asn-Gly-Thr-Gly-Lys-Asp-

                   85                  90                  95                 100
Trp-Ala-Leu-Ile-Lys-Leu-Ala-Gln-Pro-Ile-Asn-Gln-Pro-Thr-Leu-Lys-Ile-Ala-Thr-Thr-

                  105                 110                 115                 120
Thr-Ala-Tyr-Asn-Gln-Gly-Thr-Phe-Thr-Val-Ala-Gly-Trp-Gly-Ala-Asn-Arg-Glu-Gly-Gly-
                                                    ↑3
                  125                 130                 135                 140
Ser-Gln-Gln-Arg-Tyr-Leu-Leu-Lys-Ala-Asn-Val-Pro-Phe-Val-Ser-Asp-Ala-Ala-Cys-Arg-
                            ↑4
                  145                 150                 155                 160
Ser-Ala-Tyr-Gly-Asn-Glu-Leu-Val-Ala-Asn-Glu-Glu-Ile-Cys-Ala-Gly-Tyr-Pro-Asp-Thr-

                  165                 170                 175                 180
Gly-Gly-Val-Asp-Thr-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Met-Phe-Arg-Lys-Asp-Asn-Ala-

                  185                 190                 195                 200
Asp-Glu-Trp-Ile-Gln-Val-Gly-Ile-Val-Ser-Trp-Gly-Tyr-Gly-Cys-Ala-Arg-Pro-Gly-Tyr-

                  205                 210                 215                 220
Pro-Gly-Val-Tyr-Thr-Glu-Val-Ser-Thr-Phe-Ala-Ser-Ala-Ile-Ala-Ser-Ala-Ala-Arg-Thr-Leu
```

Table XIX), was a molecule which was complete, with the exception of four areas containing single residue overlaps. These areas are shown in Table XX and indicated with arrows marked 1,2,3, and 4. The attempts to gain unequivocal proof of structure about these areas, resulted in the following studies.

E. Automated 'Edman-Begg' Sequence Analysis of S-β-Aminoethylated Cyanogen Bromide Fragment Cn-1

As stated in the previous section, Table XX indicates the assumed protein sequence as established by the preceding research. The entire sequence was established with the exception of the designated single residue overlaps which were considered inadequate proof of primary sequence. However, the credibility of this sequence as established with these single residue overlaps is increased on consideration of both the known amino acid composition and the fact that in no case had a peptide been isolated which could not be placed in this sequence. The amino acid analysis previously published by Jurasek et al (1969) was used, with the exception that the number of leucines was assumed to be twelve, in agreement with the tentative sequence data (109). All residues, with the exception of aspartic acid, glutamic acid and serine, were accounted for in the 221 residue sequence by this calculation. Amino acid analysis

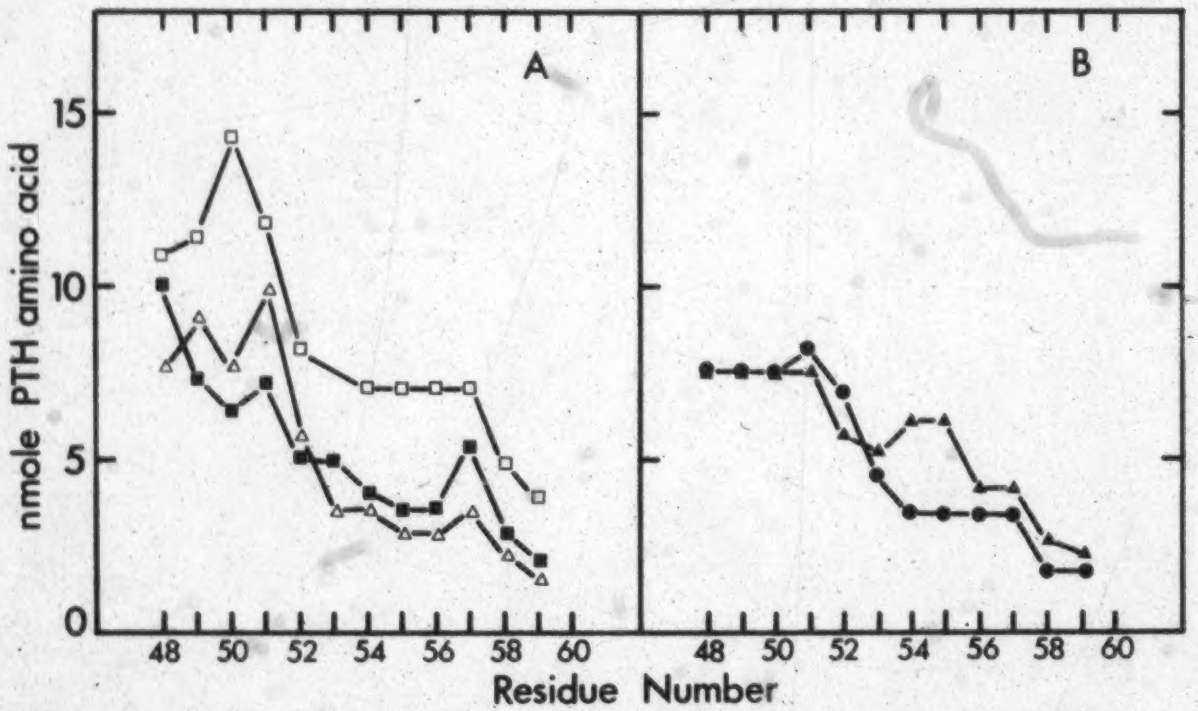


Fig. 41: Composite plots of yields for six PTH-amino acids from steps 28 to 39 in the automated sequential degradation of the cyanogen bromide fragment Cn-1. These degradation cycles correspond to residues 48 to 59 in SGT. Figures 41 A and 41 B each contain superimposed plots of three different amino acids. Figure 41 A. □——□ ,Ala; △——△ ,Thr; ■——■ ,Leu/Ile. Figure 41 B. ▲——▲ ,Val; ●——● ,Gly. Examination of the yields of each residue at each cycle allows for a qualitative estimation of the residues at those positions. Isoleucine and leucine were determined together in Fig. 41A,as yields were too low to permit differentiation of silylated derivatives.

TABLE XXI (cont'd)

Cycle	Residue Number	Residue Identified	% Yield PTH Amino Acid
23	43	Gly	7.1
24	44	Asn	1.6
25	45	Asx[d]	4.2
26	46	Thr	3.9
27	47	—	—
28	48	Leu/Ile	3.5

[a] AECys appears with dehydroserine on GLC and was differentiated by amino acid analysis of hydrolyzates.

[b] Leu and Ile differentiated by GLC of silylated derivatives and by amino acid analysis of hydrolyzates.

[c] Detected qualitatively by Pauly procedure (113).

[d] Value given assumes 100% conversion of PTH-Asx to Asp by 6 M HCl at 110° for 24 hrs in the absence of O_2.

TABLE XXI

RESULTS OF AUTOMATED SEQUENCE ANALYSIS OF 310 NANOMOLES OF S-β-AMINOETHYLATED CYANOGEN BROMIDE FRAGMENT Cn-1

Cycle	Residue Number	Residue Identified	% Yield PTH Amino Acid
1	21	Gly	5.5
2	22	AECys[a]	7.4
3	23	Gly	16.1
4	24	Gly	9.7
5	25	Ala	21.0
6	26	Leu[b]	21.6
7	27	Tyr	16.1
8	28	Ala	16.7
9	29	Gln	5.2
10	30	Asp	14.5
11	31	Ile[b]	12.6
12	32	Val	14.2
13	33	Leu[b]	11.6
14	34	Thr	10.0
15	35	Ala	11.6
16	36	Ala	11.9
17	37	His[c]	—
18	38	AECys[a]	4.5
19	39	Val	9.4
20	40	Ser	5.8
21	41	Gly	7.7
22	42	Ser	5.8

estimated the three aberrant amino acids to have values corresponding to approximately one extra residue.

In an attempt to verify the -Thr-Ala-Thr- sequence assumed from the single residue overlap designated as number 1 in Table XX, the S-β-aminoethylated Cn-1 fragment was subjected to the automated 'Edman-Begg' procedure for sequential degradation of proteins on the Beckman 890 B Sequencer, as detailed in Chapter II, Methods section N. Sequencing was thus initiated at position 21 in the molecule, which corresponds to the NH_2-terminal glycine of the Cn-1 fragment, and proceeded unambiguously, to residue 46 (see Table XXI) at which point the yields of each step became too low for identification of residues without careful consideration of the background peaks. A consideration of six PTH-amino acids from steps 28 to 39 which correspond to residues 48 to 59 in the enzyme, is shown in Figure 41. Although these data only allow for a qualitative estimation of the residue at these positions, it is of considerable significance that they agree with the sequence as assumed in Table XX. Thus, examination of these data would indicate that the sequence corresponding to the region of the molecule between residues 48 and 57 might be (Ile/Leu)-Thr-Ala-Thr-X-X-Val-Val-X-(Ile/Leu) which agrees well with the expected sequence. It should be pointed out at this time, that

the single residue overlap number 2 is at position 56 and the next residue should be a leucine if this overlap is found to be correct. It is therefore, gratifying to observe the (Ile/Leu) peak at this position in the sequential degradation, even though the result does not offer positive confirmation of this overlap. The (Ile/Leu) yields were too low to permit identification of the silylated derivatives and thus distinguish between these residues. Positions where, due to low yields, no residue can be identified are indicated by X.

The unexpectedly short sequence analysis described above, was attributed to the low solubility of the fragment. The peptide was found to be insoluble in heptafluorobutyric acid used in the cyclization step and in Quadrol buffer used in the coupling step. Indeed, the peptide was only found to be soluble in anhydrous trifluoroacetic acid, in which it was dissolved in order to transfer the sample to the reaction cup. The solubility problem with Quadrol buffer was believed to be the result of reaction of the five S-β-aminoethyl and six ε-aminolysyl groups with phenylisothiocyanate, increasing the hydrophobic character of the fragment. This would result in lowered yields throughout the sequencing procedure, due to insolubility in the buffer used in the coupling step of the Edman degradation.

In addition, another problem was uncovered on

investigation of certain peptides isolated from the α-lytic protease digest of Cn-1. As shown in Table XX the region of the molecule where yields became too low for identification of the PTH-derivatives, corresponds to a section with the sequence -Gly-Ser-Gly-Asn-Asn-Thr-Ser-Ile-Thr-. Two peptides, αI2-2 and αIII2-5, with this composition, (see Table XVIII) were isolated from the α-lytic digest. The peptide, αI2-2, was found to be neutral at pH 6.5, while αIII2-5 carried a negative charge as determined from its mobility at this pH, using the method of Offord (1966) (170). More important was the fact that the charged species could not be sequenced any further than the first asparagine, while the neutral peptide, αI2-2, showed no such apparent block. This latter peptide was obtained in 1.5% yield while the modified αIII2-5 peptide was obtained in 5.5% yield. The sequence obtained for αI2-2, using the 'dansyl-Edman' procedure, was Gly-Ser-Gly-Asn-Asn-Thr-Ser-(Ile,Thr). Since there are two asparagine residues together in the sequence, an ambiguity results in determining the actual position of the block in αIII2-5. This arose from the dansyl procedure which indicated that after the fourth and subsequent degradations, only aspartic acid was present as the NH_2-terminus. Thus, it was not possible to differentiate between a block after asparagine-44, and a block after asparagine-45. In retrospect, it is clear

now that the subtractive Edman degradation procedure would have resolved this question. However, it would appear that there is good precedence for production of a β-aspartyl peptide bond between aspartic acid and threonine but in no case is there a report of a bond of this type formed between aspartic acid and asparagine. Haley and Corcoran (1967) demonstrated that β-aspartyl-threonine and β-aspartylglycine peptides could be isolated from performic acid oxidized ribonuclease (171). All peptides were isolated from exhaustive enzymatic digests as the NH_2-terminal β-aspartyl peptides which demonstrate a typical blue ninhydrin stain. Similarly, Smyth, Stein and Moore (1962) demonstrated that a β-aspartylserine containing peptide, present in performic acid oxidized ribonuclease, was not amenable to Edman degradation beyond the β-aspartyl residue (172). The latter group suggested that this bond was formed via a cyclic imide intermediate, which was easily hydrolyzed in the alkaline pH range to form predominantly the β-aspartyl bond. Production of the imide intermediate was believed to occur during the cleavage step in the Edman degradation procedure carried out in glacial acetic acid—anhydrous HCl at 100°. This is in agreement with those conditions found by Swallow and Abraham (1959) to be ideal for formation of similar cyclic imides (173). Furthermore, Smyth et al (1963) showed in another

publication, that this particular β-aspartyl bond was not produced when anhydrous trifluoroacetic acid was used in the cleavage step, in place of the previous acid mixture (174). However, Haley and Corcoran, as mentioned previously, isolated several varieties of β-aspartylglycine peptides as well as β-aspartylthreonine from enzymic digests of performic acid oxidized ribonuclease, indicating that either the acid conditions of performic acid oxidation or those of subsequent column procedures for isolation, were instrumental in catalyzing the reaction (171).

In summary, it is believed that the conditions for cyanogen bromide fragmentation of SGT are likely responsible for the apparent deamidation and possible cyclic imide production at position 45 in this molecule. Subsequent adjustment of the pH into the alkaline range during S-β-aminoethylation could result in the β-carboxyl shift.

It was concluded that the previously mentioned solubility problem, in conjunction with the apparent β-carboxyl shift, precluded the use of the S-β-aminoethylated Cn-1 fragment from further use in the automated sequencing procedure.

It should briefly be indicated that the cause of the increasing yields over the first six steps, indicated in Table XXI, is not known, but may be due to a slow

elimination of oxygen from the reaction cup.

F. Automated 'Edman-Begg' Sequence Analysis of Native SGT

The previous attempt to provide further evidence regarding the sequence of SGT around the single residue overlap numbered 1 in Table XX, was discussed in the previous section. Two problems were elucidated; the major one involved a general reduction in yields resulting from a decreased solubility in the reaction buffer, while the other problem appeared to arise out of a β-carboxyl shift which effectively blocks the Edman degradation procedure in a high percentage of the Cn-1 peptide. In order to eliminate these problems and corroborate the result obtained in the previous section, it was decided that an attempt be made to sequence, as far as possible, the native SGT. Since the enzyme was dissolved immediately in anhydrous trifluoroacetic acid and dried down in the reaction cup of the Beckman Sequencer, no autolysis was anticipated. Similarly, the next step involving reduction of the protein with 1,4-butanedithiol in 1-chlorobutane was not expected to result in autolysis. Indeed, although activities were not measured, it was assumed that the enzyme would lose all activity as a result of the previous procedures. The denatured protein was then subjected to normal automated sequential Edman degradation as described in Chapter II,

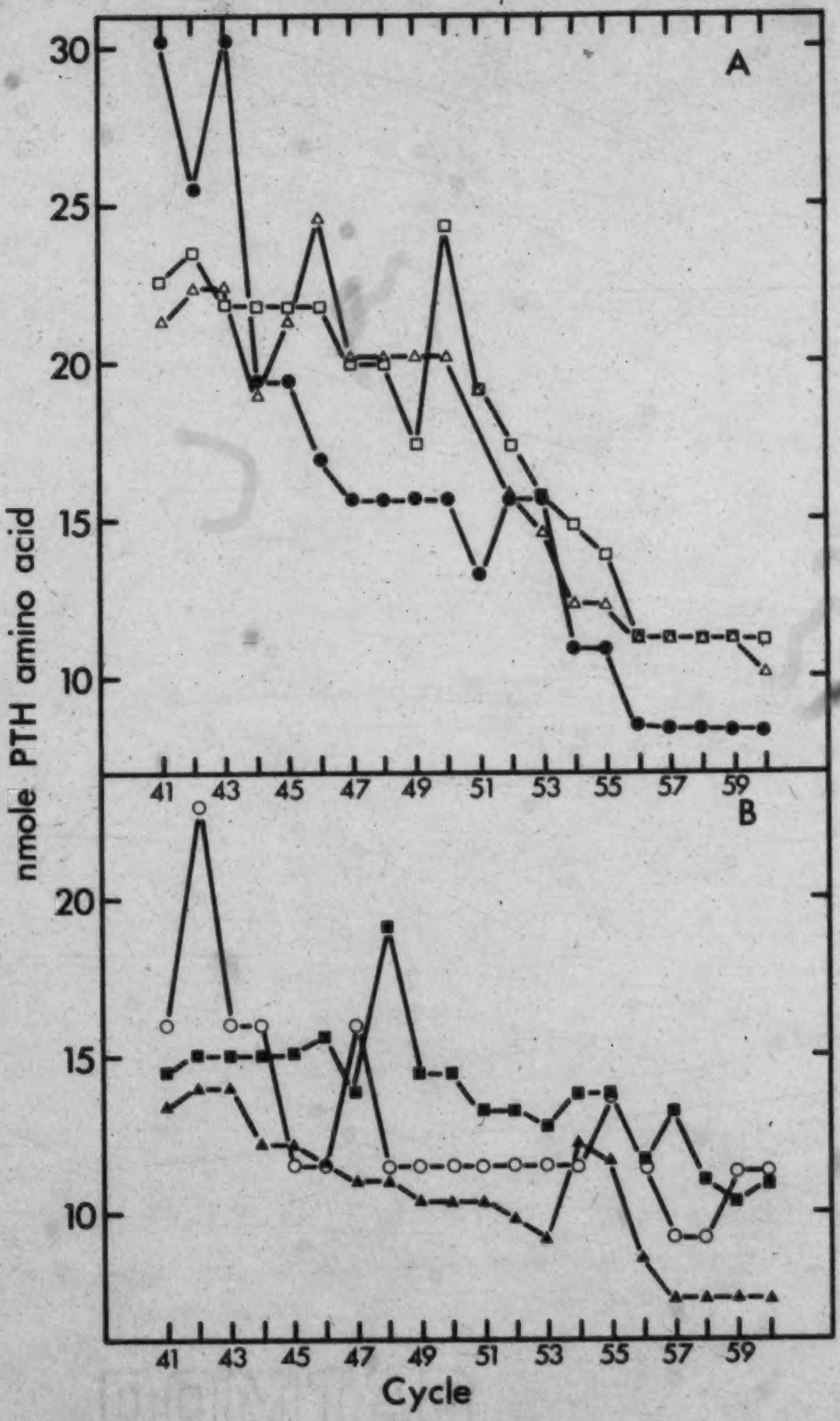

A
B
nmole PTH amino acid
Cycle
30
25
20
15
10
41
43
45
47
49
51
53
55
57
59

Figure 42

Fig. 42: See Facing Page.

Composite plots of yields for seven PTH-amino acids from steps 41 to 60 in the automated sequential degradation of SGT. Figures 42 A and 42 B each contain superimposed plots of three different amino acids.
Figure 42 A. ●——●, Gly; Δ——Δ, Thr; □——□, Ala.
Figure 42 B. ○——○, Ser; ■——■, Leu/Ile;
▲——▲ , Val.
Examination of the yields of each residue at each cycle allows for a qualitative estimation of the residues at those positions. Isoleucine and leucine were determined together, in Figure 42 B, as yields were too low to permit differentiation of silylated derivatives.

TABLE XXII

RESULTS OF AUTOMATED SEQUENCE ANALYSIS OF 450 NANOMOLES OF NATIVE SGT

Cycle	Residue Identified	% Yield PTH Amino Acid
1	Val	50.3
2	Val	44.7
3	Gly	51.0
4	Gly	37.6
5	Thr	43.6
6	——*	——
7	Ala	37.8
8	Ala	48.2
9	Gln	14.9
10	Gly	27.0
11	Glu	34.5
12	Phe	25.4
13	Pro	20.6
14	Phe	25.4
15	Met	24.0
16	Val	21.8
17	——*	——
18	Leu	14.3
19	Ser	12.0
20'	Met	16.9

* Arginine not determined.

Methods section N. The machine was programmed to perform 60 degradation steps. Table XXII indicates the % yield of the first 20 residues as ascertained by gas-phase chromatography. The cleavage products arising from the degradation steps 21 to 46 were not examined since they had previously been determined in the degradation of the Cn-1 fragment described earlier. However, the important single residue overlap region, indicated as number 1 on Table XX, was confirmed by data obtained in steps 45 to 55. This result is shown in Figure 42, where yields for seven PTH-residues are plotted for each step in the program. It is evident from these data that residues 45 to 55 include the sequence -Asn-Thr-Ser-Ile-X-Ala-X-Gly-Gly-Val-Val-. Since the sequence of the peptide αI7-3 isolated from the α-lytic digest of Cn-1 has the sequence Ala-Thr-Gly-Gly-Val-Val, there is little doubt that residue 51 is a threonine. Similarly, the peptic peptide PII-1, isolated by Dr. Jurasek, has the sequence Thr-Ser-Ile-Thr-Ala. Together these peptides prove that residues 49 and 51 are both threonines, confirming the overlap. This sequence is further corroborated by the result obtained in the previous section (see Figure 41) where automated sequence analysis of the Cn-1 fragment indicated that a threonine was present at both position 49 and 51.

It is very significant that residue 57 in this

automated degradation again appears to be (Ile/Leu) as was found in the automated sequence analysis of the Cn-1 fragment. The assumed sequence indicates that a leucine should be present in this position. Another interesting observation is that residue 59 in the sequential degradation of SGT appears to be a serine, again in agreement with the assumed sequence for the molecule. It must be reiterated however, that these residues are in low yield and only qualitative estimations can be made of results obtained after so many degradations.

G. Tryptic Digest of Performic Acid Oxidized SGT

The success of the automated sequence analysis of native SGT decreased the number of single residue overlaps to be corroborated to three and gave strong indication from two separate analyses that overlap number 2 at residue 56 was correct. In order to obtain stronger evidence for the overlap at position 56, it was decided that a different approach be attempted in order to secure a large fragment of the enzyme containing this region of the molecule. Examination of the assumed sequence indicated that trypsin should theoretically cleave the molecule at position 17, occupied by an arginine and at position 62, occupied by a lysine, producing a fragment with the desired sequence. However, the previous bad experiences with insoluble tryptic peptides, indicated

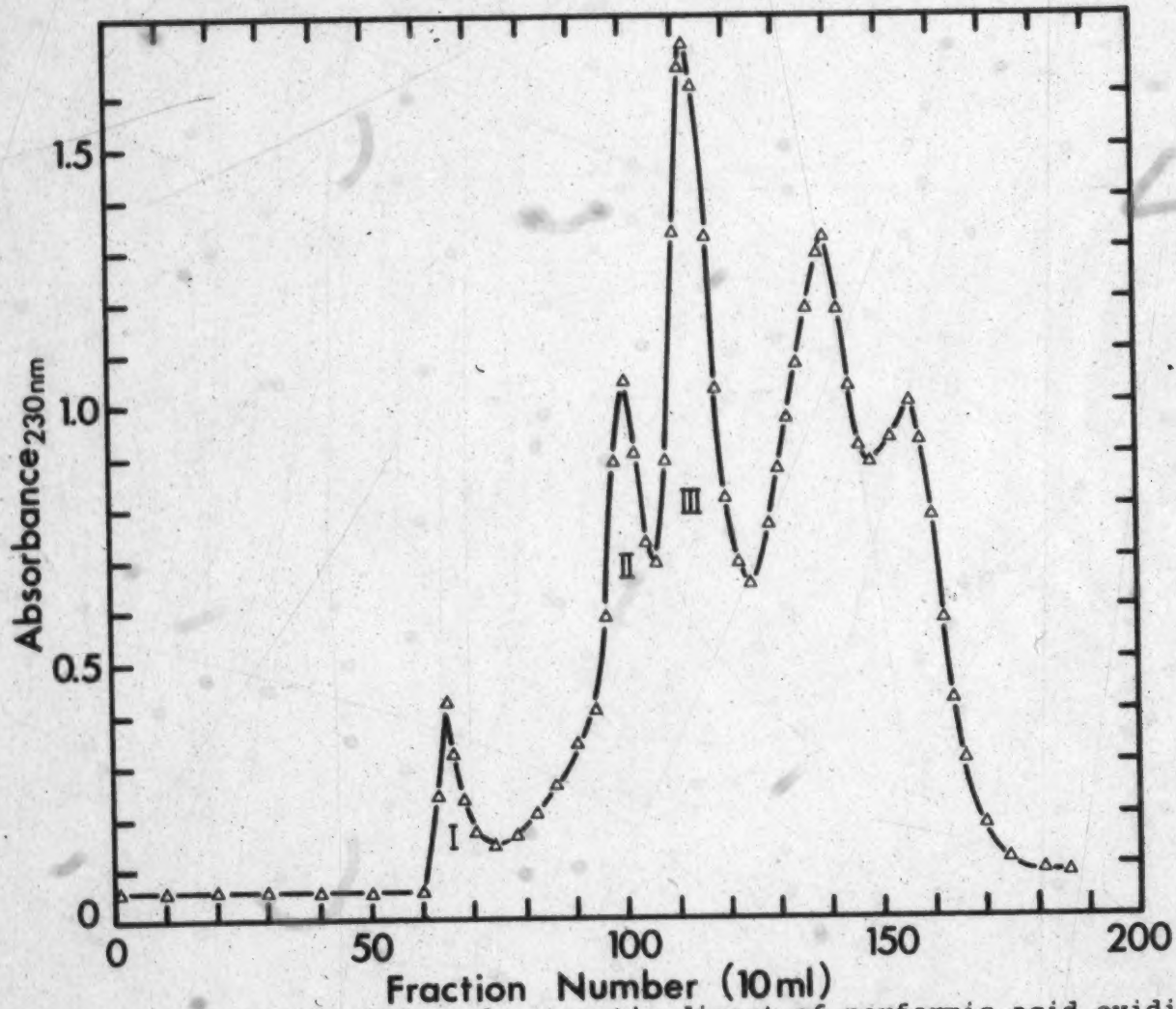


Fig. 43: Sephadex G-50 gel-filtration of a tryptic digest of performic acid oxidized SGT. A 238 mg sample was applied and eluted with 0.5M NH_4HCO_3, pH 7.8 at 60 ml/hr.

that this fragment was, in all probability, highly insoluble. This was, no doubt, due to the fact that out of the 45 residues included in this fragment only two acidic residues and two basic residues were present. Since the polypeptide contained two cysteine residues it was decided that performic acid oxidation might well increase the solubility of this fragment by doubling the number of negatively charged residues. To this end 238 mg of SGT was oxidized according to the method of Hirs (1967) as described in Chapter II, Methods section K7 and digested at a 1:100 molar ratio, (enzyme to substrate) with TPCK-trypsin as described in Chapter II, Methods section J3 (141). The reaction was terminated by freezing and lyophilization and the product dissolved in approximately 25 ml of 0.5 M NH_4HCO_3 buffer, pH 7.8. Not all of the digest mixture would dissolve in this solution, regardless of the volume used and as a result the suspension had to be clarified by centrifugation. The pellet was retained while the supernatant solution was subjected to gel filtration on a 5 x 100 cm column of Sephadex G-50 Fine equilibrated with the above buffer, as described in Chapter II, Methods section B2. The result of this chromatography is depicted in Figure 43. The second peak to elute from the column was in the theoretically predicted elution position for the fragment desired. Although the 'dansyl procedure' indicated that

leucine was the major NH_2-terminus in the peak tube (fraction 100) from peak II, this was not a satisfactory criterion in itself, of homogeneity, as at least one other tryptic split in the molecule could result in a peptide with a leucine NH_2-terminus. Furthermore, the 'dansyl procedure' can be misleading as a guide to estimation of degrees of contamination by large molecules, as in many cases the NH_2-terminal residues of large fragments do not react quantitatively, presumably as a result of partial burial of the α-amino group. The third peak to elute was expected to contain the COOH-terminal peptides, which were slightly smaller than the desired major fragment and were believed to be derived from sections of the molecule between residues 97-117, 141-176, and 178-219. The minor DNS-Ser and DNS-Bis-Lys which appeared to accompany the DNS-Leu in fraction 100 were quite possibly due to these COOH-terminal peptides. (According to results obtained by Dr. Jurasek, lysine is known to occur as the NH_2-terminal residue in a fraction of the COOH-terminal tryptic peptides as a result of hydrolysis between an -Arg-Lys- bond at positions 176 and 177 respectively.) In order to attempt a quantitation of the degree of cross-contamination of peak III with the desired material in peak II, amino acid analysis was performed on material pooled from fractions 76 to 106 inclusive. Since there was only one tyrosine in the

desired large fragment and no phenylalanines, the ratio of these residues, to the single histidine known to be present in the fragment, offered an indication of the degree of cross-contamination from extraneous peptides. The phe:his ratio indicated 36% contamination and the tyr:his ratio indicated 37% contamination. This was clearly unacceptable.

In another attempt to resolve the material in peak II from that of peak III, it was decided to attempt to run the Sephadex G-50 effluent directly into a Sephadex G-25 column. The desired leucine NH_2-terminal fragment elutes in the void volume on the latter column. It was thought possible that since peak II off the Sephadex G-50 column would enter the Sephadex G-25 bed slightly ahead of the smaller material in peak III, that the accelerated mobility on this gel system might facilitate a slightly better resolution from the following peak. Unfortunately, this was not the case; no increase in resolution was observed after this procedure. Perhaps an equilibrium between an aggregated and monomeric form of these large peptides occurred during these separation procedures.

In a last attempt to resolve the large leucine NH_2-terminal peptide from its contaminants, a 2.5 x 100 cm column containing DEAE Sephadex A-25 was equilibrated with 20 mM NH_4HCO_3 / 8 M urea at pH 9.5 and eluted with

a linear gradient system increasing from 0.02 M to 0.5 M in NH_4HCO_3 / 8 M urea. A sample containing 4.2 optical density units at 230 nm was applied and the column effluent monitored at the same wavelength. After completion of the gradient, essentially no peptide had been eluted and only a fraction of the applied sample was reclaimed on washing the column with the NH_4HCO_3 / 8 M urea solvent increased to pH 11.0 with NH_4OH. It was obvious at this time that like the previous chromatographic systems, this system could not be used to purify the peptide. Attempts at preparation of this fragment were therefore, abandoned.

H. Chymotryptic Digest of Maleylated S-β-Aminoethylated Cyanogen Bromide Fragment Cn-1

In order to prove the sequence about the remaining overlaps numbered 2 and 3 in Table XX, a method had to be devised which would facilitate solubilization of peptides produced in this area as a result of digestive procedures. Since the tryptic digestions had been shown to produce large, partially soluble, aggregating fragments, it was decided that smaller fragments might be an advantage and to this end, that a high molar ratio chymotryptic digest be attempted. Furthermore, previous experience with the solubility properties of the S-β-aminoethylated cyanogen bromide

fragment Cn-1, during the α-lytic digestion, indicated that the solubility in a basic environment was very low, and precipitation occurred with time. It was observed that the total negative charge on the Cn-1 fragment could be increased from -11 to -16 by maleylation of the S-β-aminoethylated cysteines. Indeed, when this procedure was carried out in 8 M urea as detailed in Chapter II, Methods section K3, the product produced was completely soluble after dialysis at room temperature against 50 mM N-ethylmorpholine at pH 7.5. It was this material which was subsequently used in the following chymotryptic digest.

Digestion conditions are given in detail in Chapter II, Methods section J4a, but were essentially the following: 12.5 μmoles of substrate in a 1:22 molar ratio (enzyme to substrate) were dissolved in 500 ml of 50 mM N-ethylmorpholine, pH 7.5 at 28° and reacted for 37 hours. Reaction was terminated by addition of concentrated formic acid till the pH was 3.5 and the entire solution was incubated for 6 hours at 60° to demaleylate the peptides. The reaction mixture turned opalescent shortly after the reaction was initiated and required clarification by centrifugation after the demaleylation procedure. The pellet was stored in the freezer for future use. The large volume of product was then lyophilized and finally fractionated on Dowex 1

TABLE XXIII (cont'd)

Peptide	Method of Purification	Mobility pH 6.5	Net Charge	Composition and Sequence (Values are expressed as mole ratios)	Yield µmoles
ChIII8-6	6.5	-.56	-1	(Ala,Asp,Glu,Trp) 1.01 0.95 1.03 +	0.07
ChIII9-3	6.5, 1.8	+.30	+1	(Gly,AECys,Gly,Gly,Ala,Leu,Tyr) 1.00 0.91 1.00 1.00 1.14 1.05 0.89	0.03
ChIII9-5	6.5, 1.8	-.49	-1	(Gly,Asn,Glu,Leu) 0.99 0.96 0.99 1.06	0.95
ChIII10-2	6.5, 1.8	-.15	0	(Gln,Ala,Pro,Gly,Tyr) 1.02 1.08 0.82 1.10 0.97	0.02

[a] Peptides were fractionated on Dowex 1 and Dowex 50 (peptide analyzer) prior to electrophoresis at pH(s) indicated.

[b] Mobilities were expressed relative to aspartic acid taking its value as -1.0.

[c] According to Offord (1966) (170).

[d] Glycine α-amino may have been aminoethylated and an amine may not have demaleylated.

[e] Lysine ε-amino may not have demaleylated.

TABLE XXIII (cont'd)

Peptide	Method of Purification	Mobility pH 6.5	Net Charge	Composition and Sequence (Values are expressed as mole ratios)	Yield µmoles
ChI1l-3a	6.5, 1.8	0	0	(Ala, Gln, Pro, Ile, Asn, Gln, Pro, Thr) 1.34 0.88 0.87 1.15 1.01 0.88 0.87 1.01	0.07
ChI1l-4	6.5, 1.8	-.06	0	Thr-Ala-Tyr 1.04 1.08 0.90	0.14
ChI15-9	6.5, 1.8	-.03	0	Thr-Val-Ala-Gly-Trp 0.95 0.95 1.01 1.06 +	0.44
ChIII5-4	6.5	-.57	-4	(Gly, Asn, Glu, Leu, Val, Ala, Asn, Glu, Glu, Ile, 1.12 1.03 1.10 1.10 0.93 1.10 1.03 1.10 1.10 0.98 AECys, Ala, Gly, Tyr, Pro, Asp, Thr, Gly, Gly, 0.80 1.10 1.12 0.84 0.97 1.03 0.55 1.12 1.12 Val, Asp, Thr) 0.93 1.03 0.55	0.05
ChIII7-2	6.5, 1.8	-.40	-1	(Thr, Glu, Val, Ser, Thr, Phe) 0.65 1.07 0.96 0.95 0.65 1.02	0.04
ChIII8-3a	6.5, 1.8	-.43	-2	(Glx, Ala, Pro, Gly, Tyr, Asx, Gly, Thr, Gly, Lys, 1.04 0.81 0.62 1.01 0.74 0.86 1.01 0.78 1.01 1.28 Asp, Trp) 0.86 +	0.06
ChIII8-3b	6.5, 1.8	-.43	-1[e]	(Lys, Val, Leu) 0.99 1.08 0.93	0.04

TABLE XXIII

AMINO ACID COMPOSITION AND SEQUENCE OF CHYMOTRYPTIC PEPTIDES FROM MALEYLATED CYANOGEN BROMIDE FRAGMENT Cn-1

Peptide	Method of Purification[a]	Mobility pH 6.5[b]	Net Charge[c]	Composition and Sequence (Values are expressed as mole ratios)	Yield µmoles
ChI1-1	6.5, 1.8	+.19	+1	(Ala,Gln,Pro,Ile,Asn,Gln,Pro,Thr,Leu,Lys, 1.05 1.08 0.86 0.97 0.96 1.08 0.86 0.97 0.91 0.99 Ile,Ala,Thr,Thr) 0.97 1.05 0.97 0.97	0.20
ChI1-2	6.5, 1.8	0	0	(Asn,Gln,Gly,Thr,Phe) 1.07 1.12 1.11 0.79 0.92	0.75
ChI3-8	6.5, 1.8	-.39	-1	(Ala,Gln,Asp,Ile,Val,Leu) 1.02 1.00 0.97 0.59 0.68 1.00	0.64
ChI4-7	6.5, 1.8	-.22	-1	(Glx,Ala,Pro,Gly,Tyr,Asx,Gly,Thr,Gly,Lys, 1.05 0.91 0.75 0.95 1.01 0.85 0.95 0.90 0.95 0.87 Asp,Trp) 0.85 +	0.20
ChI6-4a	6.5, 1.8	0	0[d]	(Gly,AECys,Gly,Gly,Ala,Leu,Tyr) 1.07 0.92 1.07 1.07 1.05 0.94 0.87	0.05
ChI9-5	6.5, 1.8	Insoluble at pH 6.5		(Ile,Gln,Val,Gly,Ile,Val,Ser,Trp) 1.03 0.93 0.94 1.04 1.03 0.94 0.99 +	0.12

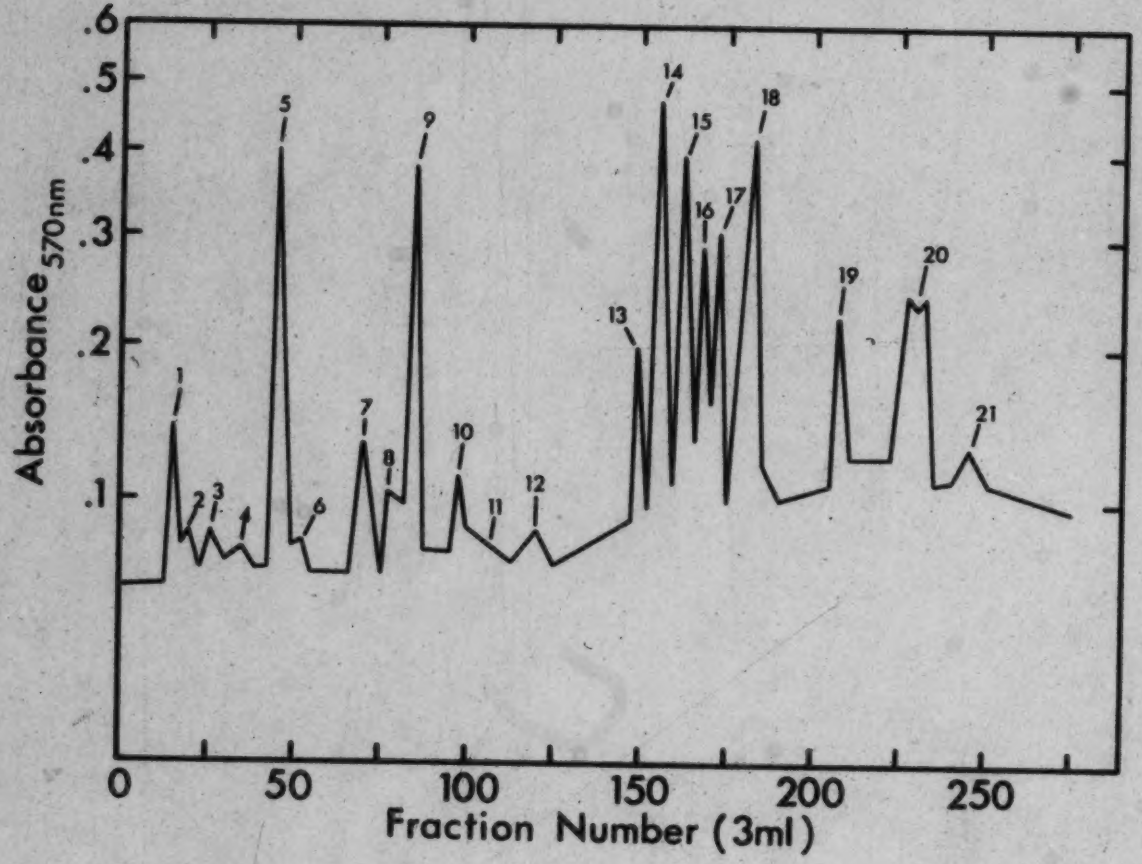


Fig. 45: Chromatography on Chromobead type P resin of fraction III from Dowex 1 fractionation of the chymotryptic digest of Cn-1. The flow rate was 25 ml/hr and 3.0 ml fractions were collected. Effluent was monitored by automatic ninhydrin analysis.

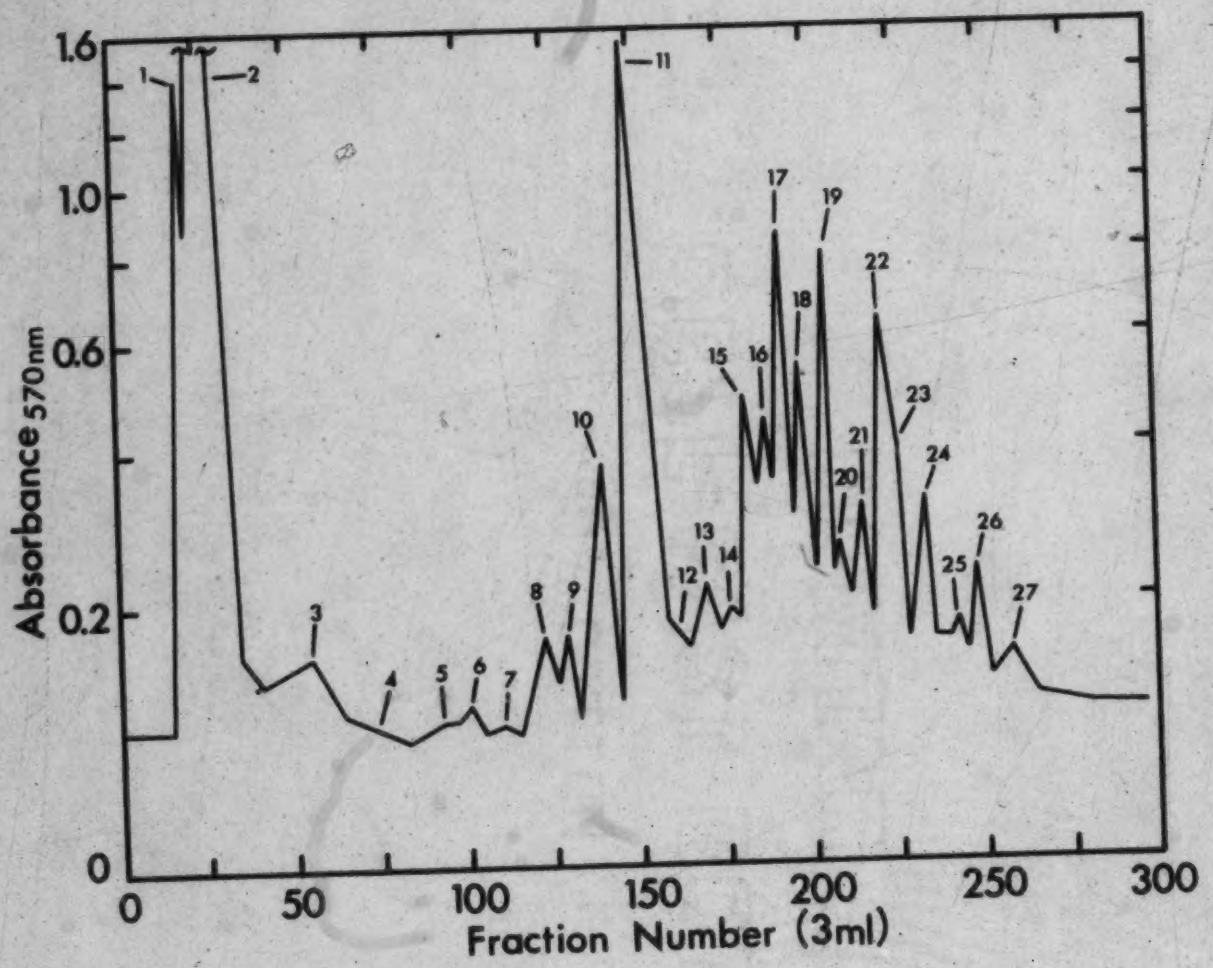


Fig. 44: Chromatography on Chromobead type P resin of fraction I from Dowex 1 fractionation of chymotryptic digest of Cn-1. The flow rate was 25 ml/hr and 3.0 ml fractions were collected. Effluent was monitored by automatic ninhydrin analysis.

as described previously for the α-lytic digest and in Chapter II, Methods section B3. The three fractions were lyophilized as before, and each further fractionated on a column of Chromobead P resin using the Technicon autoanalyzer system for peptide chromatography described in Chapter II, Methods section B4. Figures 44 and 45 indicate the profiles obtained from the basic and acidic fractions from Dowex 1 during this fractionation procedure. The neutral fraction provided a small number of peptides which were already represented in the other fractions. Again, as with the α-lytic digest peptides, the peak fractions were pooled, lyophilized, and finally subjected to electrophoresis as indicated in Table XXIII. This Table is a compilation of the data obtained during purification procedures and subsequent amino acid analyses of all the major peptides isolated.

All areas of the Cn-1 fragment are represented in this group of peptides with one exception. Unfortunately, the region required for proof of the sequence about the single residue overlap at residue 56 (number 2 in Table XX) was missing and presumed insoluble. This appeared to be the region enclosed by residues 34 to 66 inclusive. However, a tryptophan positive chymotryptic peptide, ChI15-9, was isolated in 3.5% yield and when sequenced, using the 'dansyl-Edman' procedure, provided the peptide Thr-Val-Ala-Gly-Trp.

This result corroborated the sequence associated with residue 113 or single residue overlap number 3 in Table XX. This result left only two single residue overlap sequences unproven. These regions are indicated as overlap number 2 and number 4 at positions 56 and 128 in the assumed sequence.

I. Chymotryptic Digest of Maleylated Performic Acid Oxidized SGT

While the single residue overlap number 2 appeared to be quite intractable, due to its position in a region of low charge density and due to a complete lack of useful marker residues which would facilitate isolation, the overlap number 4 appeared to offer an opportunity for isolation and separation using the lysine-diagonal technique suggested by Butler *et al* (1969) (138). It was considered likely that the peptide Leu-Leu-Lys-Ala-Asn-Val-Pro-Phe could be obtained from a chymotryptic digest if the single residue overlap at the lysine residue, indicated the correct sequence. To this end, 2.7 μmoles of SGT was oxidized in performic acid according to the method of Hirs (1967) as detailed in Chapter II, Methods section K7 (141). This step was taken to completely inhibit the enzyme and at the same time, unfold the protein by breaking the disulphide bonds, thus providing for a more accessible substrate molecule. Maleylation of the oxidized product was performed

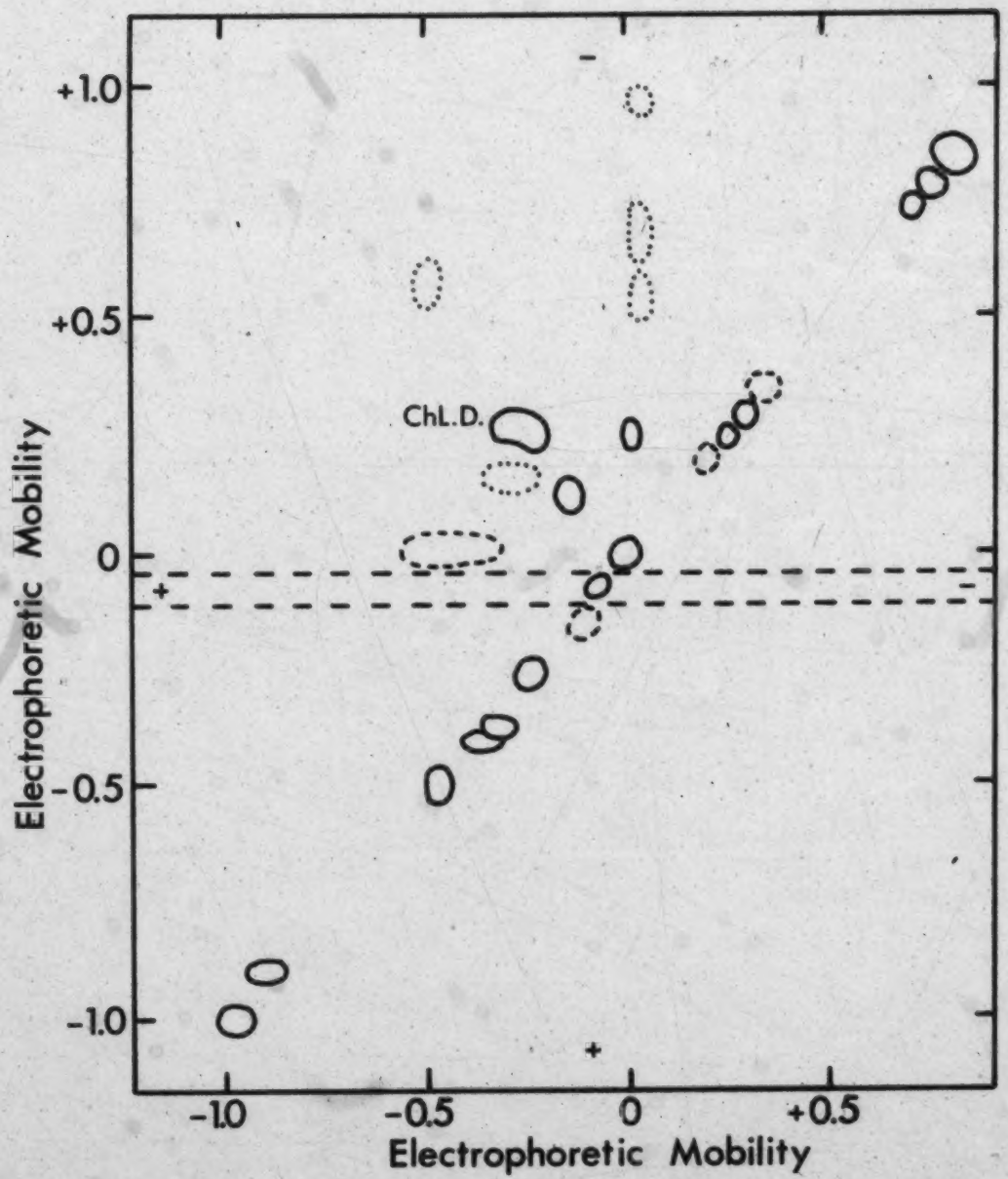


Fig. 46: A pH 6.5/pH 6.5 lysine diagonal peptide map of a chymotryptic digest of performic acid oxidized SGT. ChL.D. indicates the position of the required demaleylated lysine containing peptide described in the text.

according to the procedure of Butler et al (1969) in order that the net negative charge of the peptide be increased by two (138). The change in charge follows from the fact that maleic anhydride reacts with lysine ε-amino groups, effectively cancelling a positive charge and adding a negative charge. The substrate molecule was then dialyzed and digested at a 1:200 molar ratio (enzyme to substrate) at pH 8.0 for five hours. After applying the reaction product to a 27 cm strip of Whatman 3 MM paper, electrophoresis at pH 6.5 was performed at 60 volts/cm for 45 min and guide strips removed for cadmium-ninhydrin staining. A further side strip was removed for demaleylation in the presence of pH 3.5 pyridine-acetic acid buffer as described in Chapter II, Methods section K3. After careful drying, the side strip was sewn onto a sheet of Whatman 3 MM paper, perpendicular to the previous direction of electrophoresis and again, subjected to electrophoresis at pH 6.5. The entire sheet was finally stained with cadmium-ninhydrin and the position of the basic peptides migrating off the diagonal was noted. Since the mobility of the assumed peptide relative to aspartic acid, could be estimated according to the procedure of Offord (1966) (m'= +.25) a strip was removed from the original electropherogram that corresponded to the ninhydrin-positive spot indicated on Figure 46 as ChL.D. (Chromotryptic lysine diagonal) (170).

Accordingly, this strip was demaleylated in an identical manner to the first and the peptides separated at pH 6.5 as before. Guide strips were removed and the single basic peptide eluted from the position indicated by the cadmium-ninhydrin positive spots on the strips. Amino acid analysis of ChL.D. indicated a peptide with the following composition and residue molar ratios:

(Leu,Lys,Ala,Asn,Val,Pro,Phe).
2.02 0.90 1.13 1.04 1.03 0.84 1.01

The sequence of this peptide determined by the 'dansyl-Edman' procedure is shown below:

Leu-Leu-Lys-Ala-Asn-Val-Pro-Phe.

This result gave unambiguous proof of the sequence around the single residue overlap number 4 at residue 128 and provided the complete sequence of the molecule with the exception of that about the single residue overlap number 2 at residue 56.

As discussed earlier, the Beckman Sequencer indicated on analysis of both SGT and the Cn-1 fragment, that residue number 56 was either leucine or isoleucine and residue 59 was a serine. Although this result is in agreement with the assumed sequence in this area of the molecule, the yields of these residues do not allow for an unequivocal conclusion. However, these data serve to indicate that the primary sequence proposed is very likely correct.

CHAPTER IV

DISCUSSION

The preceding chapters have presented results of experiments directed towards elucidation of the physico-chemical properties and primary sequence of the enzyme Streptomyces griseus trypsin. Wherever possible the physico-chemical parameters measured were compared with those for bovine trypsin and other Asp-Ser-Gly serine proteases. In essentially every case the values determined for SGT compared very well with those observed for bovine trypsin. For example, stability studies carried out in the presence and absence of calcium ion, indicated that SGT was indeed more stable at higher pH values in the presence of the cation. Calcium ion is well known to have a stabilizing effect on bovine trypsin and other serine proteases (47). SGT was also found to undergo what is believed to be reversible denaturation at pH 12.0 and pH 2.0, a phenomenon previously observed with bovine trypsin (49). Another similarity with bovine trypsin is the molecular weight of SGT, estimated by ultracentrifugation studies, and corroborated by sequence analysis, to be 22,900. The molecular weight of bovine trypsin is approximately 24,000 (48). Still other results are indicative of the general similarity of SGT with the trypsin family. For example, the $K_{m(app)}$ for SGT, using BAEE as substrate, was found to be approximately

8×10^{-6} M, a value very similar to 10^{-5} M observed by Gutfreund for bovine trypsin (150). Similarly, the dependence of enzyme activity on pH in the acidic range was found to depend upon a basic group with an apparent pK_a of 6.25 for bovine trypsin and 5.84 for SGT. In both α-chymotrypsin and trypsin this group was found to be one of the histidine residues (53). Since SGT has only one histidine residue the $pK_{(app)}$ observed must be that of the imidazole of histidine 57, assuming that the same general mechanism of catalysis occurs in SGT as in the other serine proteases. Several lines of evidence strongly support this assumption. For example, SGT is specifically inhibited by the 'active-site-directed' inhibitor TLCK, in an analogous manner to that observed by Mares-Guia and Shaw (1963) for trypsin (31). Indeed, the modified active site histidine was isolated and demonstrated to be alkylated after inhibition with this reagent. Furthermore, this irreversible inhibition could be blocked by the reagent p-aminobenzamidine shown by Mares-Guia et al (1965), to be a strong competitive inhibitor of bovine trypsin (158). It is also known that SGT is sensitive to DFP, with the inhibited enzyme being phosphorylated at a single serine residue (16). Subsequent isolation of a phosphorylated peptide with the sequence Asp-Ser-Gly clearly demonstrated that the enzyme belonged to the

Asp-Ser-Gly group of serine proteases and no doubt catalyzed reactions with the same basic charge transfer mechanism for all members of this group (16). SGT was also shown to catalyze the hydrolysis of BAEE at a rate approximately 1/3 that of normal, when deuterium oxide was used in place of water in the reaction mixture. This too, has been demonstrated to be characteristic of this class of esterases, although the precise interpretation of the result may be questionable (161). The preceding results indicate that the Asp-Ser-Gly serine protease SGT, has a catalytic mechanism that is not unlike bovine trypsin and other members of this group of proteases.

An extensive specificity study using oxidized insulin A and B chains, has indicated that SGT like other trypsins is specific for arginine and lysine residues, in agreement with results published earlier from this laboratory (109). It was noted, during these experiments, however, that all but the most highly purified SGT was contaminated with trace amounts of a 'non-trypsin-like' activity. This activity could be partially resolved from the major SGT peak on a long Bio-Rex-70 column. In retrospect, it is now clear that an affinity chromatography system as utilized by Robinson et al (1971) for isolation of bovine, porcine and dogfish trypsin, would have been very useful in the final purification

of SGT for the specificity studies (162).

All of the preceding results have strongly supported the tenet that SGT is physically and enzymatically very similar to bovine trypsin. One group of studies, directed towards elucidation of the role of the NH_2-terminal valine α-amino group in SGT, was less successful in making a clear correlation with the known information for bovine trypsin. As discussed in both Chapter I and Chapter III, α-chymotrypsin and trypsin are known to be inactivated when their α-amino groups are chemically modified (78,79,163). In both enzymes the loss in activity paralleled the modification of the NH_2-terminal α-amino group. In the case of α-chymotrypsin only isoleucine-16 was found to obey this relationship, while the other two NH_2-terminal α-amino groups reacted independently of the loss in activity (163). Kinetic studies of the pH dependence of enzyme activity, as well as spectropolarimetric investigations, have implicated the isoleucine-16 NH_2-terminal α-amino group of α-chymotrypsin in the maintenance of substrate binding ability and an active conformation (68,69). This result was corroborated by the X-ray diffraction studies of α-chymotrypsin by Blow et al (1969), which showed that the isoleucine-16 α-amino group forms an ion pair with aspartic acid-194 in the active site (82). Stroud et al (1971) have similarly shown that the trypsin NH_2-terminus

forms a salt bridge in the active site of that protease (35). Not all serine proteases however, demonstrate a requirement for such a group. For example, α-lytic protease does not appear to have a requirement for a charged NH_2-terminal α-amino group for activity and exhibits no conformational changes on titration in the alkaline pH range (75,76).

The NH_2-terminal studies on SGT were conducted at the time when most of the previously mentioned studies were being reported, and resulted in essentially two conclusions. It was observed that the NH_2-terminal α-amino group was inaccessable for chemical modification unless denatured, in agreement with the findings of Robinson et al (1973) for bovine trypsin (79). Secondly, there appeared to be no group with an apparent pK_a of 8 to 9 which was required for the maintenance of an active enzyme conformation. The latter conclusion follows as a result of titration studies which indicated that SGT activity does not begin to decrease till approximately pH 10.0. Furthermore, ORD and CD studies conducted at various pH values throughout the alkaline range showed no pH dependent conformational change until the enzyme began to denature at pH values greater than 10.0. Thus, there is no evidence from the results obtained to implicate the role of an ionized α-amino group in the activity and conformational

stabilization of SGT. Indeed, if such a charged group is required it only becomes available for titration on general unfolding and denaturation of the enzyme. This conclusion is in general agreement with that made by Shotton and Watson (1970) for elastase, in their review of the structural and enzymatic properties of that enzyme (34). It is interesting to note that the first five residues of porcine elastase are identical with those of SGT.

The latter conclusion was the only result obtained during the investigation of SGT that indicated a marked physico-chemical difference from bovine trypsin. However, it should be noted that attempts to modify the NH_2-terminal of SGT were greatly hampered by the lack of an inactive precursor, and as a result, no absolute decision regarding the role of the NH_2-terminal α-amino group should be made until an adequate chemical modification experiment is performed. In spite of the high reactivity of nitrous acid, the deamination procedure as employed by Hofmann (1967) might be the most expedient means of modifying a protease which lacks a known zymogen (78).

The structural studies directed towards elucidation of the primary sequence of SGT have provided the most unequivocal evidence of the similarity of SGT and bovine trypsin. An early report from this laboratory

(Jurasek et al (1969)) stressed the remarkable degree of homology existing between the disulphide bridge peptides of these two enzymes and indicated that, even from this preliminary study, the Streptomyces enzyme appeared to be intermediate in structure between bovine trypsin and α-lytic protease of Myxobacter 495 (109). This conclusion was drawn from the observation that, like α-lytic protease, SGT had only three disulphide bridges linking corresponding regions of primary structure, and only a single histidine. On the other hand, the specificity and disulphide bridge peptides of SGT indicated a remarkable similarity with bovine trypsin. This conclusion has been amply corroborated by the elucidation of the complete sequence of SGT. For example, the percentage of identical residues between SGT and bovine trypsin, on alignment for maximal homology, was found to be 34%. When conservative replacements are counted, this homology increases to 41%. (Conservative replacements were scored allowing for the following very restrictive replacements: Asn=Gln, Ile=Leu=Val, Ala=Gly, Thr=Ser, Lys=Arg, Tyr=Phe=Trp, Asp=Glu.) In contrast, α-lytic protease has only 21% identity and 32% homology on comparison of its sequence with that of the assumed SGT sequence. Clearly, on a basis of sequence homology, SGT appears to be most similar to bovine trypsin. However, as mentioned previously, the

fact that SGT has only three disulphide bridges and only one histidine, in contrast with six disulphide bridges and three histidines in bovine trypsin, would indicate that SGT is evolutionarily intermediate between the latter enzyme and α-lytic protease which is similar, in these respects, to SGT. It is interesting to note that the assumed SGT sequence shows a 32% identity with the sequence of bovine α-chymotrypsin—remarkably similar to the 34% identity seen with trypsin. This may indicate that both α-chymotrypsin and trypsin have diverged from a common ancestor with SGT, at approximately the same time.

In the general introduction, the active site sequences of SGT were compared with those of other serine proteases to indicate the general similarity of these enzymes. In the following pages other highly homologous areas of SGT will be compared with other serine proteases and their positions indicated on a ribbon diagram of the B- and C-chains of α-chymotrypsin. The activation peptide was eliminated for clarity and the arrangement of polypeptide chain and homologous side chains was made identical to that in chymotrypsin. Insertions and deletions in SGT were ignored during this initial construction. It was obvious from examination of the α-chymotrypsin model that the differences in amino acid sequences could be readily accommodated without any

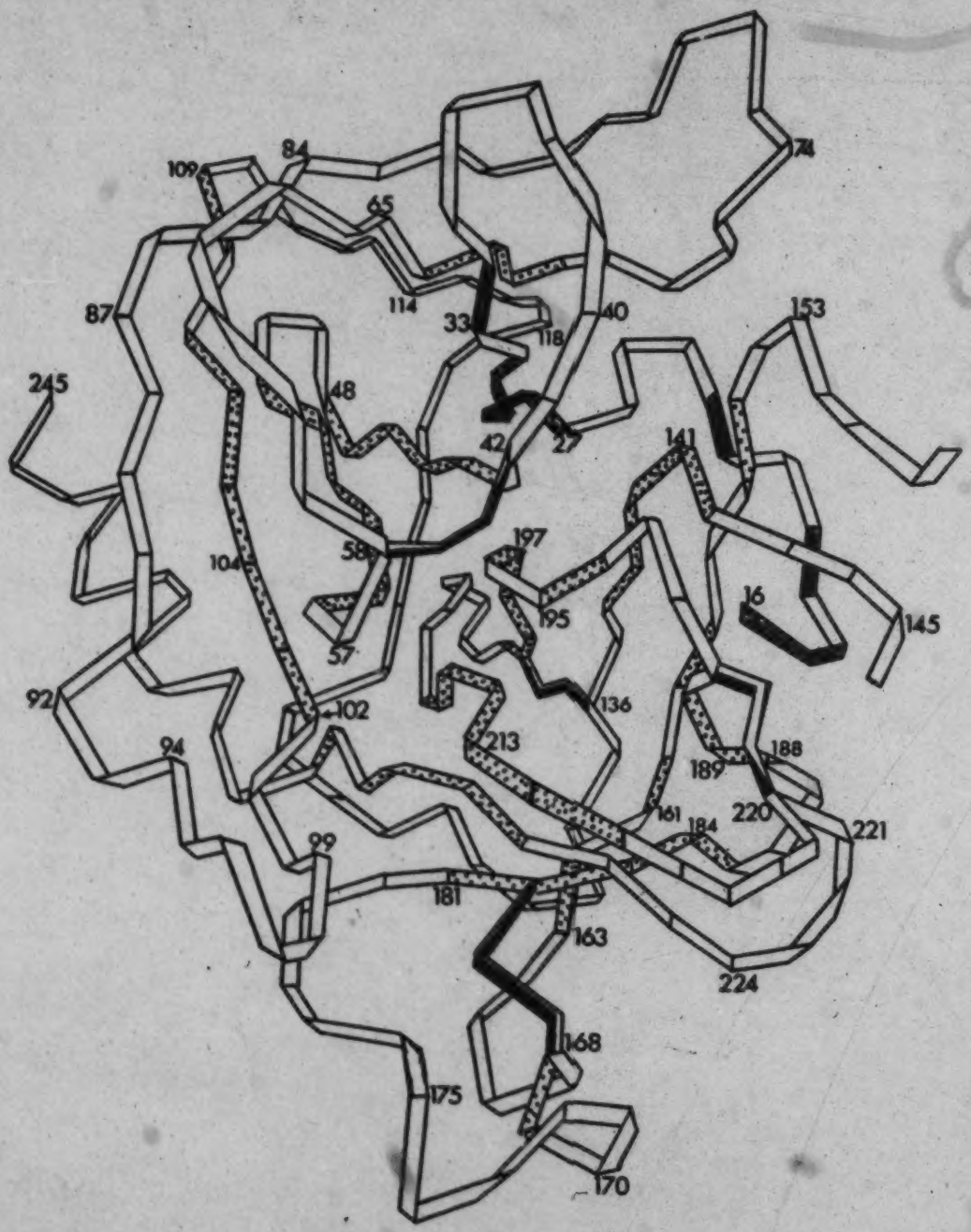


Fig. 47: Ribbon diagram of B- and C-chains of α-chymotrypsin with homologous residues of SGT from Table XXIV indicated as shaded bars. Buried residues are stippled.

TABLE XXIV

HOMOLOGOUS SEQUENCES ADJACENT TO THE NH_2-TERMINI OF α-LYTIC PROTEASE (α-LP), PORCINE ELASTASE (PE), BOVINE CHYMOTRYPSIN A (BCA), BOVINE CHYMOTRYPSIN B (BCB), BOVINE TRYPSIN (BT) AND *STREPTOMYCES GRISEUS* TRYPSIN (SGT)

	15 A	15 B	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
α-LP:	Ala	Asn	Ile	Val	Gly	Gly	Ile	Glu	Tyr	Ser	Ile	Asn	Asn	-	-	-	-	-	-	-	-
PE:			Val	Val	Gly	Gly	Thr	Glu	Ala	Gln	Arg	Asn	Ser	Trp	Pro	Ser	Gln	Ile	Ser	Leu	Gln
BCA:			Ile	Val	Asn	Gly	Glu	Glu	Ala	Val	Pro	Gly	Ser	Trp	Pro	Trp	Gln	Val	Ser	Leu	Gln
BCB:			Ile	Val	Asn	Gly	Glu	Asp	Ala	Val	Pro	Gly	Ser	Trp	Pro	Trp	Gln	Val	Ser	Leu	Gln
BT:			Ile	Val	Gly	Gly	Tyr	Thr	Cys	Gly	Ala	Asn	Thr	Val	Pro	Tyr	Gln	Val	Ser	Leu	Asn
SGT:			Val	Val	Gly	Gly	Thr	Arg	Ala	Ala	Gln	Gly	Glu	Phe	Pro	Phe	Met	Val	Arg	Leu	Ser

serious steric problems. Indeed, in no case did a hydrophilic group in the SGT sequence replace an internal hydrophobic group of chymotrypsin. In order to expand on the similarities and differences observed on construction of this ribbon model, the previously mentioned active site sequences, as well as the other areas of homology, will be placed in the diagram.

Beginning at the NH_2-terminus of SGT, Table XXIV indicates homologous sequences in this region of the molecule as compared with bovine trypsin, chymotrypsins A and B, porcine elastase and α-lytic protease of Myxobacter 495. Figure 47 indicates the positions of these homologous residues on the previously mentioned ribbon diagram of chymotrypsin. They are indicated as shaded sections and to facilitate the spacial comprehension of the diagram, the parts of sequence which are buried under the surface of the molecule are indicated as stippled sections. The numbering system used throughout, is that of α-chymotrypsin. It is perhaps significant that the residues proximal to the NH_2-terminus of these serine proteases have a high degree of homology. Indeed, this might be anticipated if the proposed role of the NH_2-terminal α-amino group, in stabilizing the active conformation of the molecule via an ion pair with aspartic acid-194, is to be a common property of this group of proteases.

Fig. 50: Ribbon diagram of B- and C-chains of α-chymotrypsin with homologous residues of SGT from Table XXVII added to those shown in Fig. 49 and indicated as shaded bars.

TABLE XXVII

HOMOLOGOUS SEQUENCES IN THE REGION OF THE DISULPHIDE BRIDGE RESIDUES HALF CYSTINE 168 AND 182

	167	168	169	170							180	181	182	183
*														
α-LP:	Val	Cys	Arg	Ser	-	-	-	-	-	-	-		Cys	Gly
PE:	Ile	Cys	Ser	Ser	-	-	-	-	-		-Met	Val	Cys	Ala
BCA:	Asn	Cys	Lys	Lys	-	-	-	-	-		-Met	Ile	Cys	Ala
BCB:	Asp	Cys	Arg	Lys	-	-	-	-	-		-Met	Ile	Cys	Ala
BT:	Ser	Cys	Lys	Ser	-	-	-	-	-		-Met	Phe	Cys	Ala
SGT:	Ala	Cys	Arg	Ser	-	-	-	-	-		-Glu	Ile	Cys	Ala

* See Table XXIV for enzyme abbreviations.

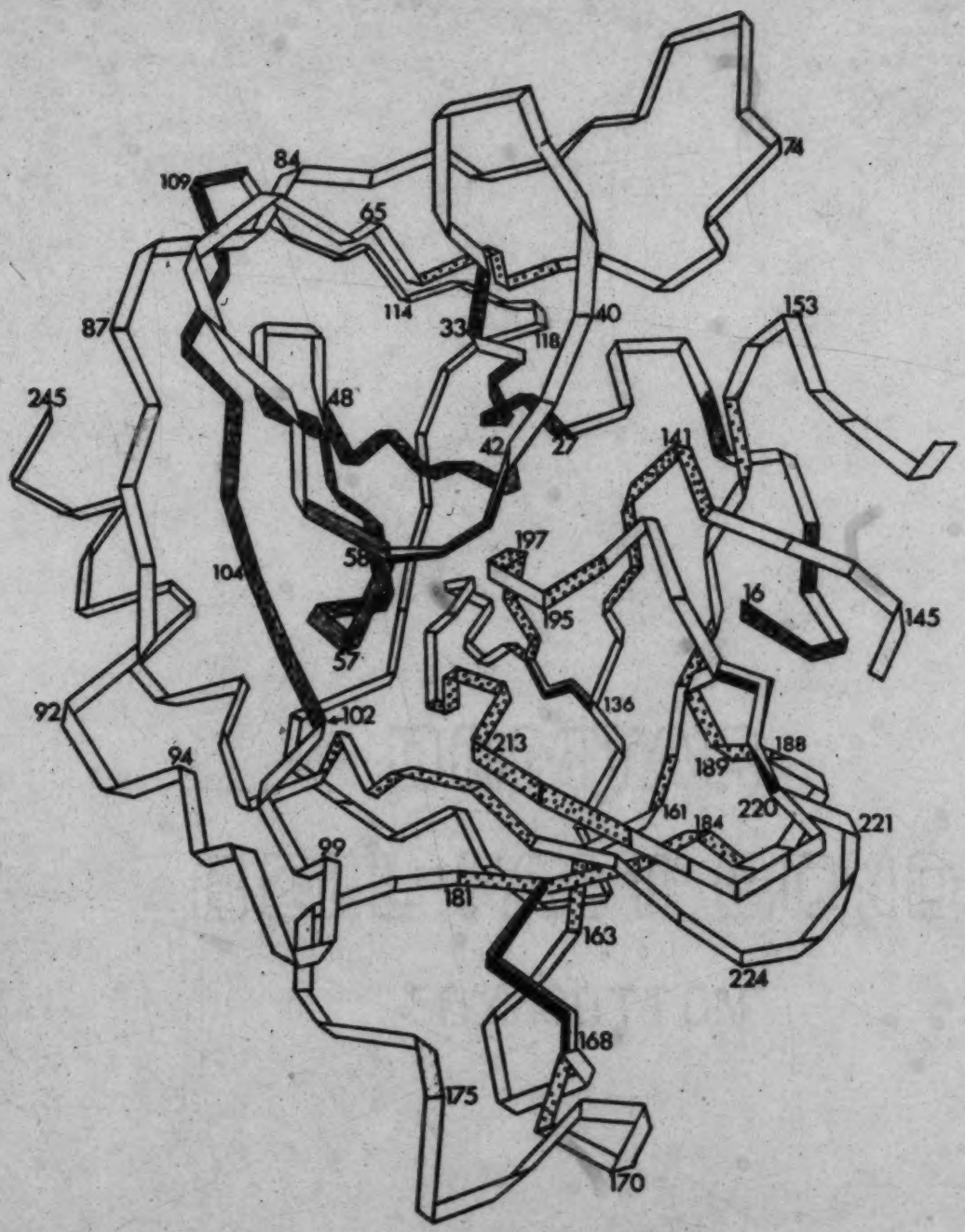


Fig. 49: Ribbon diagram of B- and C-chains of α-chymotrypsin with homologous residues of SGT from Table XXVI added to those shown in Fig. 48 and indicated as shaded bars.

TABLE XXVI

HOMOLOGOUS SEQUENCES IN THE REGION OF ASPARTIC ACID-102

	100	101	102	103	104	105	106	107	108	109
SGPA*:	Asn	Asp	Asp	Tyr	Gly	Ile	Ile	Arg	His	Ser
α-LP:	Gly	Asn	Asp	Arg	Ala	Trp	Val	Ser	Leu	Thr
PE:	Gly	Tyr	Asp	Ile	Ala	Leu	Leu	Arg	Leu	Ala
BCA:	Asn	Asn	Asp	Ile	Thr	Leu	Leu	Lys	Leu	Ser
BCB:	Arg	Asn	Asp	Ile	Thr	Leu	Leu	Lys	Leu	Ala
BT:	Asn	Asn	Asp	Ile	Met	Leu	Ile	Lys	Leu	Lys
SGT:	Gly	Lys	Asp	Trp	Ala	Leu	Ile	Lys	Leu	Ala

* *Streptomyces griseus* protease A (SGPA); for other enzyme abbreviations see Table XXIV.

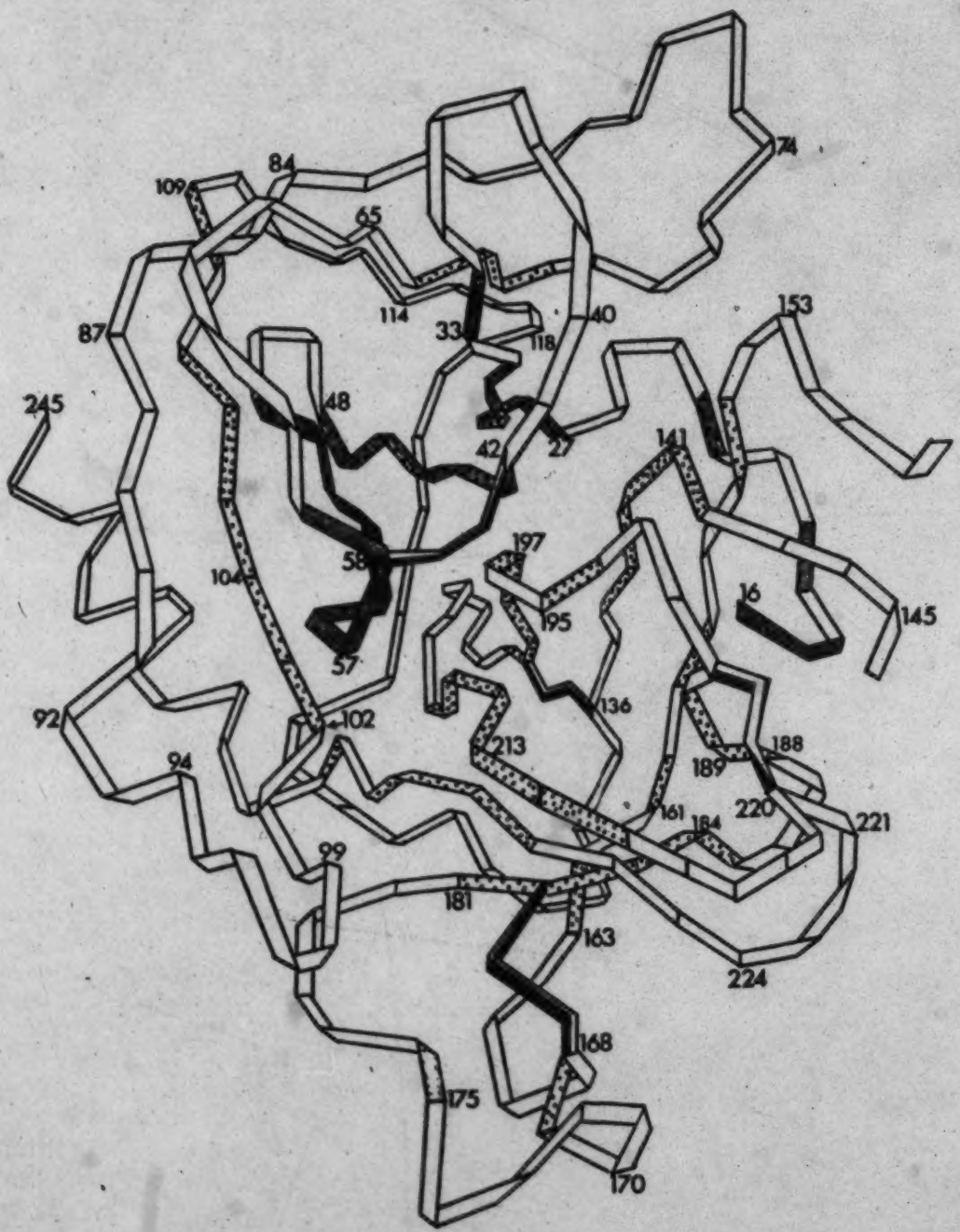


Fig. 48: Ribbon diagram of B- and C-chains of α-chymotrypsin with homologous residues of SGT from Table XXV added to those shown in Fig. 47 and indicated as shaded bars.

TABLE XXV

HOMOLOGOUS SEQUENCES IN THE REGION OF THE DISULPHIDE BRIDGE RESIDUES HALF CYSTINE 42 AND 58

	40	41	42	42 A	43	44	44 A	45	46	47				51	52	53	54	55	56	57	58	59
SGPA*:	Ser	Arg	Cys	Ser	Leu	Gly	Phe	Asn	Val	Ser	-	-	-	His	Ala	Leu	Thr	Ala	Gly	His	Cys	Thr
α-LP:	Ser	Leu	Cys	Ser	Val	Gly	Phe	Ser	Val	Thr	-	-	-	Gly	Phe	Val	Thr	Ala	Gly	His	Cys	Gly
PE:	His	Thr	Cys	-	Gly	Gly	-	Thr	Leu	Ile	-	-	-	Trp	Val	Met	Thr	Ala	Ala	His	Cys	Val
BCA:	His	Phe	Cys	-	Gly	Gly	-	Ser	Leu	Ile	-	-	-	Trp	Val	Val	Thr	Ala	Ala	His	Cys	Gly
BCB:	His	Phe	Cys	-	Gly	Gly	-	Ser	Leu	Ile	-	-	-	Trp	Val	Val	Thr	Ala	Ala	His	Cys	Gly
BT:	His	Phe	Cys	-	Gly	Gly	-	Ser	Leu	Ile	-	-	-	Trp	Val	Val	Ser	Ala	Ala	His	Cys	Tyr
SGT:	Met	Gly	Cys	-	Gly	Gly	-	Ala	Leu	Tyr	-	-	-	Ile	Val	Leu	Thr	Ala	Ala	His	Cys	Val

* Streptomyces griseus protease A (SGPA); for other enzyme abbreviations see Table XXIV.

Table XXV indicates a comparison of homologous sequences about the cysteine residue 42 and its disulphide bridge mate, cysteine-58. The latter residue is adjacent to the reactive histidine residue number 57. As described earlier, an extensive homology can be observed near the histidine-57 residue. It should be noted that sequence from another enzyme, Streptomyces griseus protease A (SGPA), has been added to the top of this table. Again, these homologous areas are shown on the diagram, in addition to the previously mentioned homologies. (see Figure 48) Similarly, the highly homologous aspartic acid-102 sequence is shown in Table XXVI and the corresponding position of these residues shown in Figure 49. It is also interesting to note that SGT contains a tyrosine at position 94. A similar tyrosine found in this position in chymotrypsin is on the surface of the molecule and has been suggested to be responsible in part, for burying aspartic acid-102 in the active site of chymotrypsin (82). The conservation of this residue in the primary sequence of SGT lends support to this proposal.

Further homology near the disulphide bridge connecting cysteine-168 and 182 is presented in Table XXVII. The positions of the homologous residues are again added to Figure 50 along with the other areas of homology. Blow and Steitz (1970) have suggested that

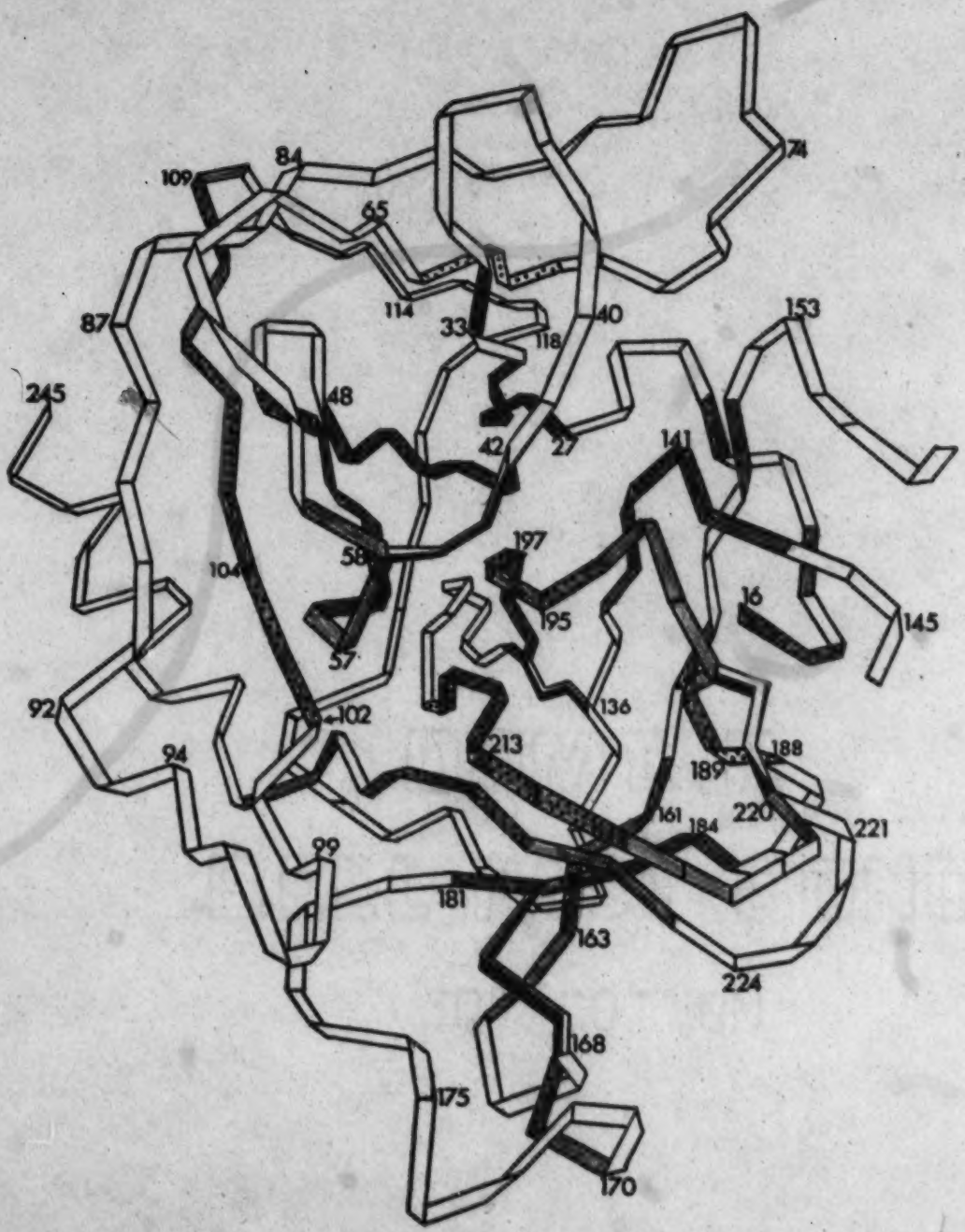


Fig. 52: Ribbon diagram of B- and G-chains of α-chymotrypsin with homologous residues of SGT from Table XXX added to those shown in Fig. 51 and indicated as shaded bars.

TABLE XXX

HOMOLOGOUS SEQUENCES IN THE REGION OF THE COOH-TERMINI OF SOME SERINE PROTEASES

	224	225	226	227	228	229	230	231	232	233	234	235
* α-LP:	Arg	Ser	Ser	Leu	Phe	Glu	Arg	Leu	Gln	Pro	Ile	Leu
PE:	Lys	Pro	Thr	Val	Phe	Thr	Arg	Val	Ser	Ala	Tyr	Ile
BCA:	Thr	Pro	Gly	Val	Tyr	Ala	Arg	Val	Thr	Ala	Leu	Val
BCB:	Thr	Pro	Ala	Val	Tyr	Ala	Arg	Val	Thr	Ala	Leu	Met
BT:	Lys	Pro	Gly	Val	Tyr	Thr	Lys	Val	Cys	Asn	Tyr	Val
SGT:	Tyr	Pro	Gly	Val	Tyr	Thr	Glu	Val	Ser	Thr	Phe	Ala

* See Table XXIV for enzyme abbreviations.

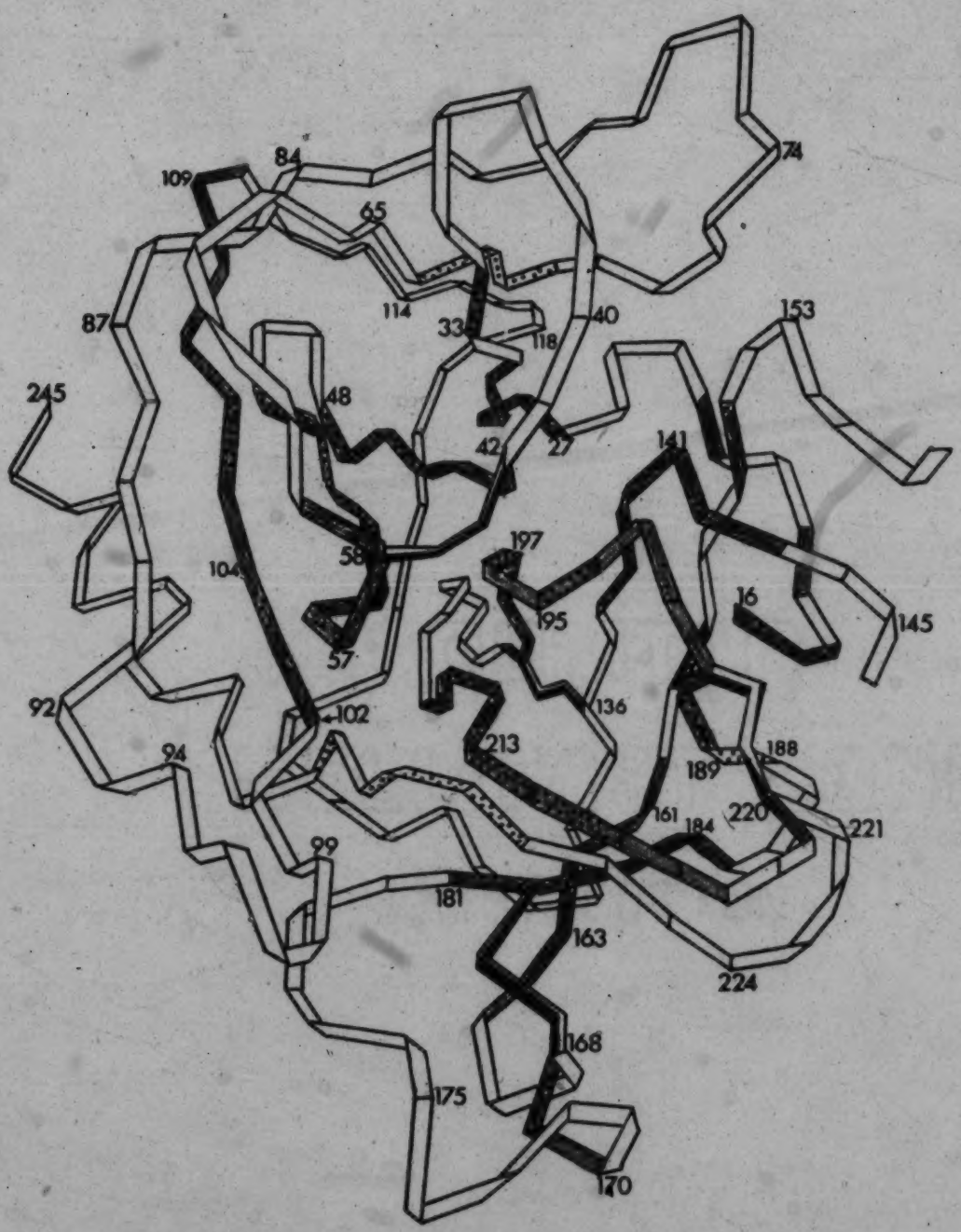


Fig. 51: Ribbon diagram of B- and C-chains of α-chymotrypsin with homologous residues of SGT from Tables XXVIII and XXIX added to those shown in Fig. 50 and indicated as shaded bars.

TABLE XXIX

HOMOLOGOUS SEQUENCES ASSOCIATED WITH THE SPECIFICITY POCKET IN TWO REGIONS OF THE SERINE PROTEASES

| | 137 | 138 | 139 | 140 | 141 | 142 | | |

TABLE XXVIII

HOMOLOGOUS SEQUENCES IN THE REGION OF THE ACTIVE SERINE-195

	189	189 A	189 B	189 C	189 D	189 E	190	191	192	192 A	192 B	193	194	195	196	197	198	199	200	201
SGPA*:	Gly	Met	Ile	Gln	Thr	Asn	Val	Cys	Ala	Gln	Pro	Gly	Asp	Ser	Gly	Gly	Ser	Leu	Phe	Ala
α-LP:	Gly	Leu	Thr	Gln	Gly	Asn	Ala	Cys	Met	Gly	Arg	Gly	Asp	Ser	Gly	Gly	Ser	Trp	Ile	Thr
PE:	Ser	-	-	-	-	-	Gly	Cys	Gln	-	-	Gly	Asp	Ser	Gly	Gly	Pro	Leu	His	Cys
BCA:	Ser	-	-	-	-	-	Ser	Cys	Met	-	-	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Val	Cys
BCB:	Ser	-	-	-	-	-	Ser	Cys	Met	-	-	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Val	Cys
BT:	Asp	-	-	-	-	-	Ser	Cys	Gln	-	-	Gly	Asp	Ser	Gly	Gly	Pro	Val	Val	Cys
SGT:	Asp	-	-	-	-	-	Thr	Cys	Gln	-	-	Gly	Asp	Ser	Gly	Gly	Pro	Met	Phe	Arg

* Streptomyces griseus protease A (SGPA); for other enzyme abbreviations see Table XXIV.

this disulphide bridge serves to stabilize the methionine loop in chymotrypsin (164). If this is true, conservation of this disulphide bridge is important in proteases where an unfolded chain would be readily autolyzed. Table XXVIII demonstrates the homology near the reactive serine-195 residue discussed in the general introduction. Together with aspartic acid-102 and histidine-57, these residues form the charge relay system. The charge relay proceeds from aspartic acid-102 to histidine-57 and eventually increases the nucleophilicity of the serine-195 hydroxyl oxygen (32). It should be indicated that aspartic acid-189, in the above sequences, is found in the specificity site of bovine trypsin and is responsible for binding the basic side chains of trypsin substrates (35). It is very significant therefore, that SGT also has an aspartic acid in this position in the molecule.
Table XXIX indicates two regions of the molecule associated with the general architecture of the specificity pocket, in which aspartic acid-189 is located. The high degree of homology observed is presumably related to this structural function. Figure 51 shows where the homologous residues from Tables XXVIII and XXIX are added to the ribbon diagram.

The last table (Table XXX) shows the homologous area found near the COOH-terminal of SGT. Figure 52 again, shows the position of these residues on the

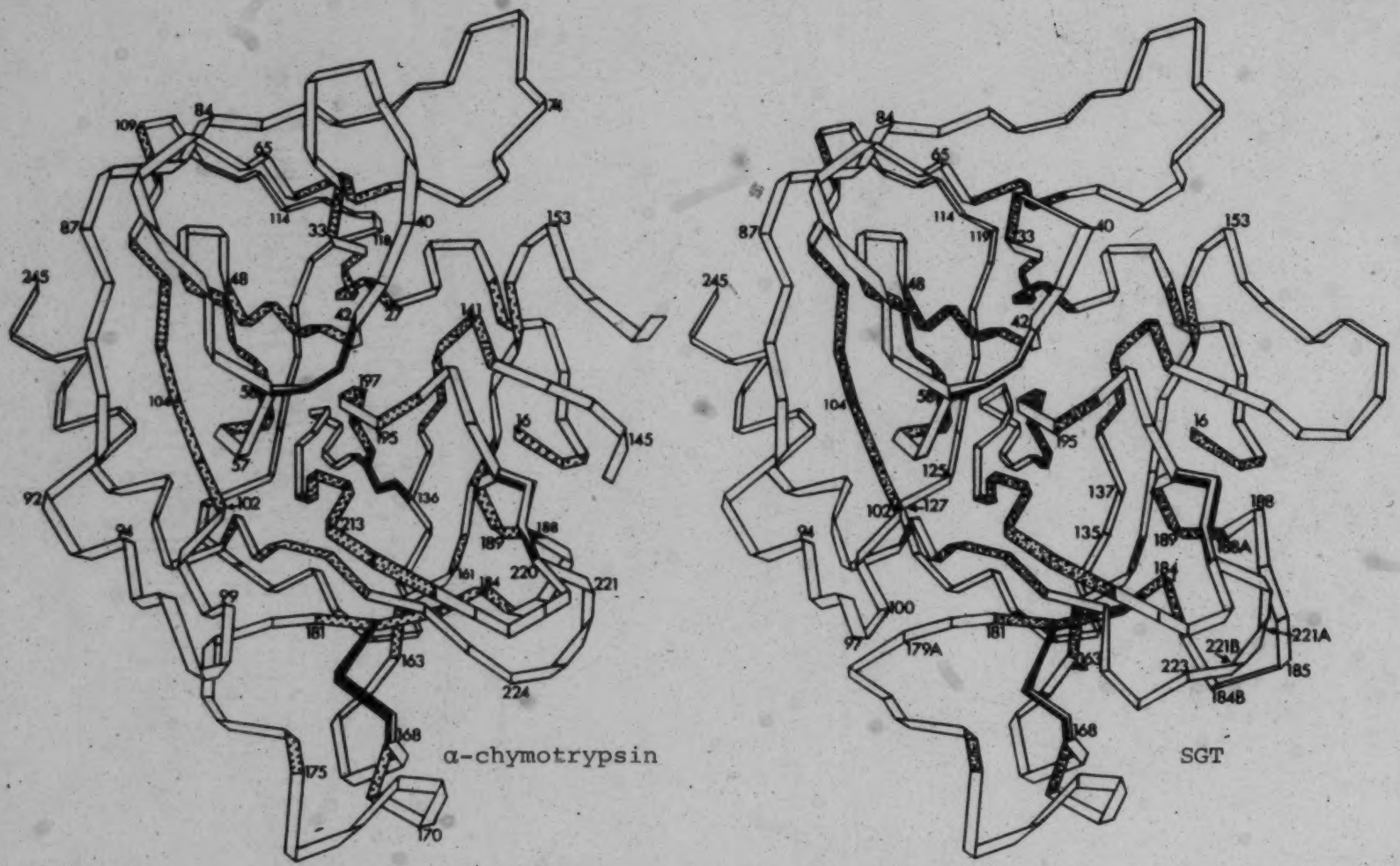


Fig. 53: A schematic diagram of chymotrypsin compared with a speculative three dimensional diagram of SGT, corrected for deletions and insertions.

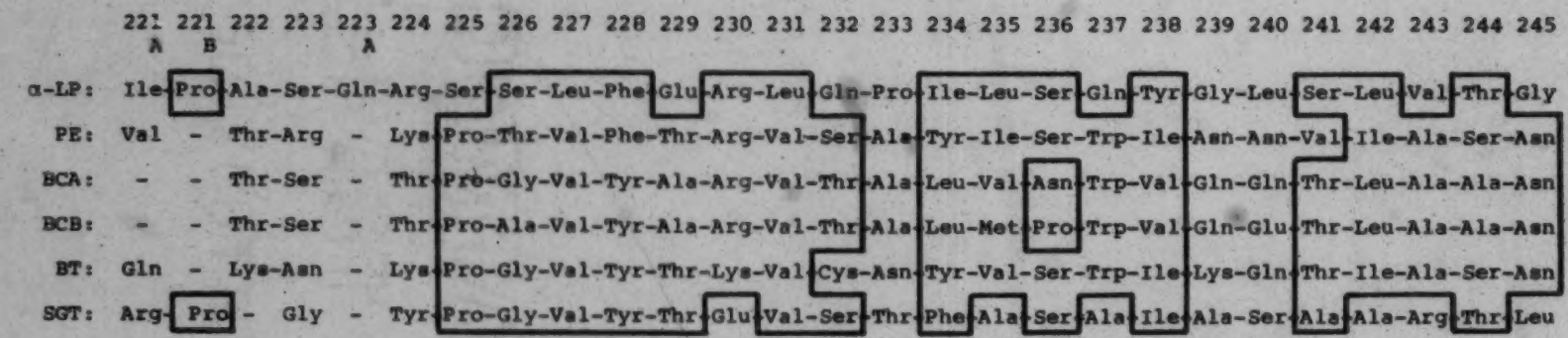
theoretical diagram. It should be noticed by now, that the homologous areas indicated by shading, overlap most of the buried areas, which are stippled in this diagram. This is an important observation, since it shows that it would be possible to build a Kendrew model of SGT having the core similar to that of chymotrypsin. Almost all residues known to be important for the enzyme's action are attached to this core. Therefore, placing these residues in the right configuration on this theoretical model of SGT would provide no difficulties. An even more useful reference molecule would, of course, be bovine trypsin. However, the co-ordinates for the DIP-trypsin structure described by Stroud et al (1971) have not been published to date (35). As a result, a speculative three dimensional model of SGT, constructed using chymotrypsin as a reference molecule, is compared with a schematic model of chymotrypsin in Figure 53. In this schematic diagram, all deletions and insertions are accounted for. The easiest detectable difference involves the disulphide bridges; there are three in SGT, five in chymotrypsin and six in trypsin. Two other major differences involve deletions; one deletion occurs in the loop encompassed by residues 102 to 125, while another is observed in the surface loop near the active site (the loop encompassed by residues 32 to 42). On the other hand, one of the important loops associated with

construction of the specificity pocket is enlarged by the insertion of residues at positions 184A, 184B and 188A.

The results presented in this analysis of the theoretical tertiary structure of SGT support the hypothesis that this enzyme has a similar three-dimensional structure to α-chymotrypsin and most likely the other serine proteases. Furthermore, it would appear that these results lend further support to the tenet that the structure of this enzyme is intermediate between bovine trypsin and α-lytic protease from Myxobacter 495.

The core of the molecule, along with the active site, seems to be the most conservative part of the molecule and was apparently able to withstand the constant mutational pressure lasting since bacteria and mammals evolved from a common ancestor. On the other hand, extensive changes appear permissable on the surface of the molecule and as a result of this, large areas of the sequence of SGT differ beyond recognition from the mammalian serine proteases. This is in agreement with the study done by Stroud et al (1971) who found that only 10% of the solvent-accessible side chains were identical in bovine trypsin, bovine chymotrypsin and porcine elastase (35).

I would like now to finally compare some of the findings discussed by Stroud et al (1971), as a consequence


221 A 221 B 222 223 223 A 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245
α-LP: Ile Pro Ala Ser Gln Arg Ser Ser Leu Phe Glu Arg Leu Gln Pro Ile Leu Ser Gln Tyr Gly Leu Ser Leu Val Thr Gly
PE: Val - Thr Arg - Lys Pro Thr Val Phe Thr Arg Val Ser Ala Tyr Ile Ser Trp Ile Asn Asn Val Ile Ala Ser Asn
BCA: - - Thr Ser - Thr Pro Gly Val Tyr Ala Arg Val Thr Ala Leu Val Asn Trp Val Gln Gln Thr Leu Ala Ala Asn
BCB: - - Thr Ser - Thr Pro Ala Val Tyr Ala Arg Val Thr Ala Leu Met Pro Trp Val Gln Glu Thr Leu Ala Ala Asn
BT: Gln - Lys Asn - Lys Pro Gly Val Tyr Thr Lys Val Cys Asn Tyr Val Ser Trp Ile Lys Gln Thr Ile Ala Ser Asn
SGT: Arg Pro - Gly - Tyr Pro Gly Val Tyr Thr Glu Val Ser Thr Phe Ala Ser Ala Ile Ala Ser Ala Ala Arg Thr Leu

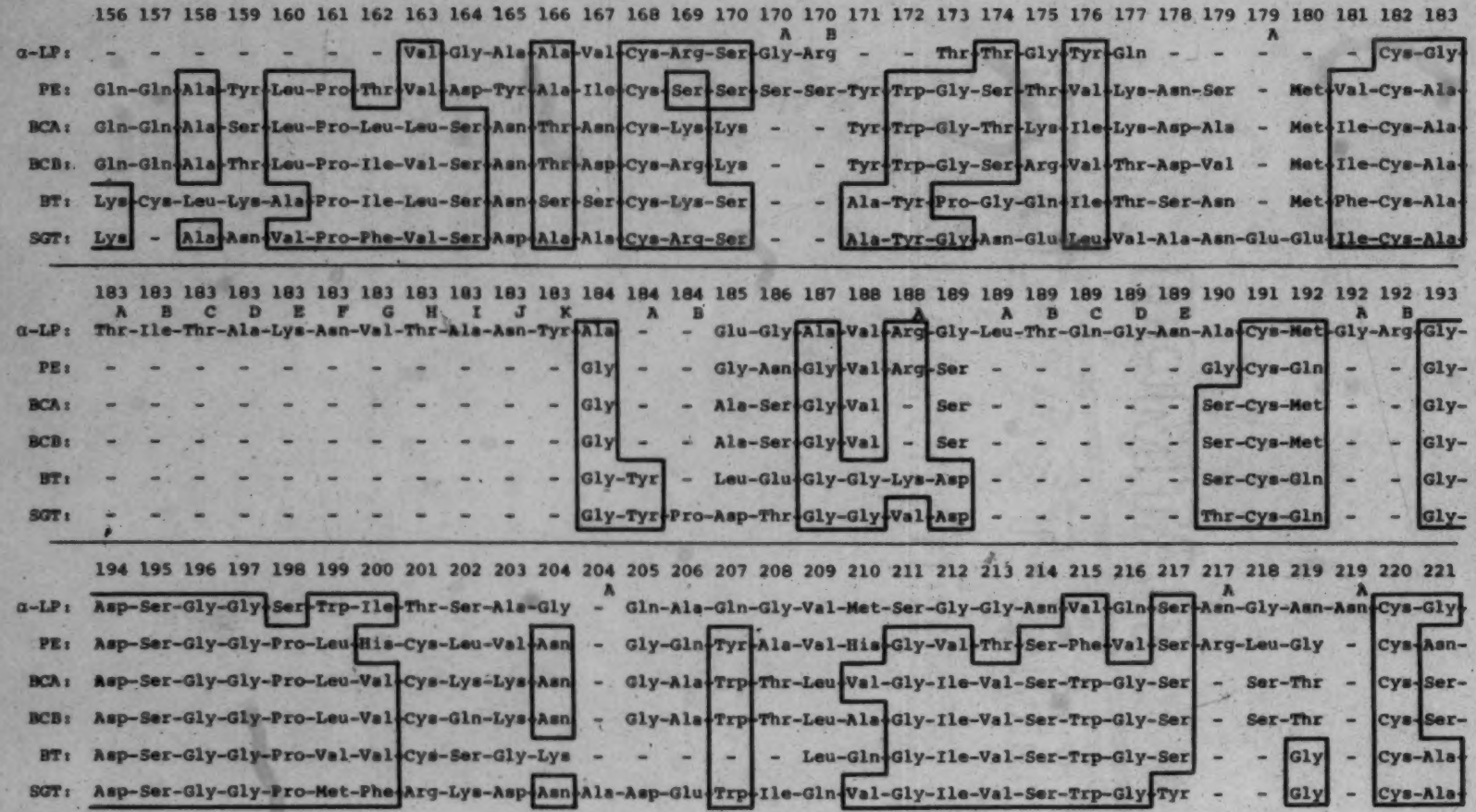

156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 170A 170B 171 172 173 174 175 176 177 178 179 179A 180 181 182 183
α-LP: - - - - - - - Val-Gly-Ala-Ala-Val-Cys-Arg-Ser-Gly-Arg - - Thr-Thr-Gly-Tyr-Gln - - - - - Cys-Gly-
PE: Gln-Gln-Ala-Tyr-Leu-Pro-Thr-Val-Asp-Tyr-Ala-Ile-Cys-Ser-Ser-Ser-Ser-Tyr-Trp-Gly-Ser-Thr-Val-Lys-Asn-Ser - Met-Val-Cys-Ala-
BCA: Gln-Gln-Ala-Ser-Leu-Pro-Leu-Leu-Ser-Asn-Thr-Asn-Cys-Lys-Lys - - Tyr-Trp-Gly-Thr-Lys-Ile-Lys-Asp-Ala - Met-Ile-Cys-Ala-
BCB: Gln-Gln-Ala-Thr-Leu-Pro-Ile-Val-Ser-Asn-Thr-Asp-Cys-Arg-Lys - - Tyr-Trp-Gly-Ser-Arg-Val-Thr-Asp-Val - Met-Ile-Cys-Ala-
BT: Lys-Cys-Leu-Lys-Ala-Pro-Ile-Leu-Ser-Asn-Ser-Ser-Cys-Lys-Ser - - Ala-Tyr-Pro-Gly-Gln-Ile-Thr-Ser-Asn - Met-Phe-Cys-Ala-
SGT: Lys - Ala-Asn-Val-Pro-Phe-Val-Ser-Asp-Ala-Ala-Cys-Arg-Ser - - Ala-Tyr-Gly-Asn-Glu-Leu-Val-Ala-Asn-Glu-Glu-Ile-Cys-Ala-
183A 183B 183C 183D 183E 183F 183G 183H 183I 183J 183K 184 184A 184B 185 186 187 188 188A 189 189A 189B 189C 189D 189E 190 191 192 192A 192B 193
α-LP: Thr-Ile-Thr-Ala-Lys-Asn-Val-Thr-Ala-Asn-Tyr-Ala - - Glu-Gly-Ala-Val-Arg-Gly-Leu-Thr-Gln-Gly-Asn-Ala-Cys-Met-Gly-Arg-Gly-
PE: - - - - - - - - - - - Gly - - Gly-Asn-Gly-Val-Arg-Ser - - - - - Gly-Cys-Gln - - Gly-
BCA: - - - - - - - - - - - Gly - - Ala-Ser-Gly-Val - Ser - - - - - Ser-Cys-Met - - Gly-
BCB: - - - - - - - - - - - Gly - - Ala-Ser-Gly-Val - Ser - - - - - Ser-Cys-Met - - Gly-
BT: - - - - - - - - - - - Gly-Tyr - Leu-Glu-Gly-Gly-Lys-Asp - - - - - Ser-Cys-Gln - - Gly-
SGT: - - - - - - - - - - - Gly-Tyr-Pro-Asp-Thr-Gly-Gly-Val-Asp - - - - - Thr-Cys-Gln - - Gly-
194 195 196 197 198 199 200 201 202 203 204 204A 205 206 207 208 209 210 211 212 213 214 215 216 217 217A 218 219 219A 220 221
α-LP: Asp-Ser-Gly-Gly-Ser-Trp-Ile-Thr-Ser-Ala-Gly - Gln-Ala-Gln-Gly-Val-Met-Ser-Gly-Gly-Asn-Val-Gln-Ser-Asn-Gly-Asn-Asn-Cys-Gly-
PE: Asp-Ser-Gly-Gly-Pro-Leu-His-Cys-Leu-Val-Asn - Gly-Gln-Tyr-Ala-Val-His-Gly-Val-Thr-Ser-Phe-Val-Ser-Arg-Leu-Gly - Cys-Asn-
BCA: Asp-Ser-Gly-Gly-Pro-Leu-Val-Cys-Lys-Lys-Asn - Gly-Ala-Trp-Thr-Leu-Val-Gly-Ile-Val-Ser-Trp-Gly-Ser - Ser-Thr - Cys-Ser-
BCB: Asp-Ser-Gly-Gly-Pro-Leu-Val-Cys-Gln-Lys-Asn - Gly-Ala-Trp-Thr-Leu-Ala-Gly-Ile-Val-Ser-Trp-Gly-Ser - Ser-Thr - Cys-Ser-
BT: Asp-Ser-Gly-Gly-Pro-Val-Val-Cys-Ser-Gly-Lys - - - - - Leu-Gln-Gly-Ile-Val-Ser-Trp-Gly-Ser - - Gly - Cys-Ala-
SGT: Asp-Ser-Gly-Gly-Pro-Met-Phe-Arg-Lys-Asp-Asn-Ala-Asp-Glu-Trp-Ile-Gln-Val-Gly-Ile-Val-Ser-Trp-Gly-Tyr - - Gly - Cys-Ala-

TABLE XXXI (cont'd)

	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	80 A	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95
α-LP:	Ile	Gly	Gly	Ala	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Val	Val	Gly	Thr	Phe	Ala	Ala
PE:	Val	Val	Val	Gly	Glu	His	Asn	Leu	Asn	Gln	Asn	Asn	Gly	Thr	Glu	-	Gln	Tyr	Val	Gly	Val	Gln	Lys	Ile	Val	Val	His	Pro	Tyr	Trp	Asn	
BCA:	Val	Val	Ala	Gly	Glu	Phe	Asp	Gln	Gly	Ser	Ser	Ser	Glu	Lys	Ile	-	Gln	Lys	Leu	Lys	Ile	Ala	Lys	Val	Phe	Lys	Asn	Ser	Lys	Tyr	Asn	
BCB:	Val	Val	Ala	Gly	Glu	Phe	Asp	Gln	Gly	Leu	Glu	Thr	Glu	Asp	Thr	-	Gln	Val	Leu	Lys	Ile	Gly	Lys	Val	Phe	Lys	Asn	Pro	Lys	Phe	Ser	
BT:	Leu	-	-	Gly	Gln	Asp	Asn	Ile	Asn	Val	Val	Glu	Gly	Asn	Gln	-	Gln	Phe	Ile	Ser	Ala	Ser	Lys	Ser	Ile	Val	His	Pro	Ser	Tyr	Asn	
SGT:	Ser	-	-	Ile	Thr	Ala	Thr	Gly	Gly	Val	Val	Asp	Leu	Gln	Ser	Ala	Val	Lys	Val	Arg	Ser	Thr	Lys	Val	Leu	Gln	Ala	Pro	Gly	Tyr	Asn	

	96	97	98	99	99 A	99 B	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124
α-LP:	Arg	-	-	Val	Phe	Pro	Gly	Asn	Asp	Arg	Ala	Trp	Val	Ser	Leu	Thr	Ser	Ala	Gln	Thr	-	-	-	-	-	-	-	-	Leu	Leu	Pro
PE:	Thr	Asp	Asp	Val	Ala	Ala	Gly	Tyr	Asp	Ile	Ala	Leu	Leu	Arg	Leu	Ala	Gln	Ser	Val	Thr	Leu	Asn	Ser	Tyr	Val	Gln	Leu	Gly	Val	Leu	Pro	
BCA:	Ser	Leu	Thr	Ile	-	-	Asn	Asn	Asp	Ile	Thr	Leu	Leu	Lys	Leu	Ser	Thr	Ala	Ala	Ser	Phe	Ser	Gln	Thr	Val	Ser	Ala	Val	Cys	Leu	Pro	
BCB:	Ile	Leu	Thr	Val	-	-	Arg	Asn	Asp	Ile	Thr	Leu	Leu	Lys	Leu	Ala	Thr	Pro	Ala	Gln	Phe	Ser	Glu	Thr	Val	Ser	Ala	Val	Cys	Leu	Pro	
BT:	Ser	Asn	Thr	Leu	-	-	Asn	Asn	Asp	Ile	Met	Leu	Ile	Lys	Leu	Lys	Ser	Ala	Ala	Ser	Leu	Asn	Ser	Arg	Val	Ala	Ser	Ile	Ser	Leu	Pro	
SGT:	Gly	Thr	-	-	-	-	Gly	Lys	Asp	Trp	Ala	Leu	Ile	Lys	Leu	Ala	Gln	Pro	Ile	Asn	Gln	-	-	-	-	Pro	Thr	Leu	Lys	Ile	Ala	

	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155
α-LP:	Arg	Val	Ala	Asn	Gly	-	-	-	Ser	Ser	Phe	Val	Thr	Val	Arg	Gly	Ser	Thr	Glu	Ala	Ala	-	-	-	-	-	-	-	-	-	-
PE:	Arg	Ala	Gly	Thr	Ile	Leu	Ala	Asn	Asn	Ser	Pro	Cys	Tyr	Ile	Thr	Gly	Trp	Gly	Leu	Thr	Arg	-	Thr	Asn	Gly	Gln	Leu	Ala	Gln	Thr	Leu	
BCA:	Ser	Ala	Ser	Asp	Asp	Phe	Ala	Ala	Gly	Thr	Thr	Cys	Val	Thr	Thr	Gly	Trp	Gly	Leu	Thr	Arg	Tyr	Thr	Asn	Ala	Asn	Thr	Pro	Asp	Arg	Leu	
BCB:	Ser	Ala	Asp	Glu	Asp	Phe	Pro	Ala	Gly	Met	Leu	Cys	Ala	Thr	Thr	Gly	Trp	Gly	Lys	Thr	Lys	Tyr	Asn	Ala	Leu	Lys	Thr	Pro	Asp	Lys	Leu	
BT:	Thr	-	Ser	Cys	Ala	Ser	-	Ala	Gly	Thr	Gln	Cys	Leu	Ile	Ser	Gly	Trp	Gly	Asn	Thr	Lys	Ser	Ser	Gly	Thr	Ser	Tyr	Pro	Asp	Val	Leu	
SGT:	Thr	-	Thr	Thr	Ala	Tyr	Asn	Gln	Gly	Thr	Phe	-	Thr	Val	Ala	Gly	Trp	Gly	Ala	Asn	Arg	Glu	Gly	Gly	Ser	Gln	Gln	Arg	Tyr	Leu	Leu	

TABLE XXXI

COMPARISON OF AMINO ACID SEQUENCES OF α-LYTIC PROTEASE (α-LP), PORCINE ELASTASE (PE), BOVINE CHYMOTRYPSINS A AND B (BCA & BCB), BOVINE TRYPSIN (BT), AND STREPTOMYCES GRISEUS TRYPSIN-LIKE ENZYME (SGT)

	15	15A	15B	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	36A	36B	36C	37	38	39	40	41
α-LP:	Ala		Asn	Ile	Val	Gly	Gly	Ile	Glu	Tyr	Ser	Ile	Asn	Asn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Ala	Ser	Leu
PE:				Val	Val	Gly	Gly	Thr	Glu	Ala	Gln	Arg	Asn	Ser	Trp	Pro	Ser	Gln	Ile	Ser	Leu	Gln	Tyr	Arg	Ser	Gly	Ser	Ser	Trp	Ala	His	Thr
BCA:				Ile	Val	Asn	Gly	Glu	Glu	Ala	Val	Pro	Gly	Ser	Trp	Pro	Trp	Gln	Val	Ser	Leu	Gln	Asp	Lys	-	-	-	Thr	Gly	Phe	His	Phe
BCB:				Ile	Val	Asn	Gly	Glu	Asp	Ala	Val	Pro	Gly	Ser	Trp	Pro	Trp	Gln	Val	Ser	Leu	Gln	Asp	Ser	-	-	-	Thr	Gly	Phe	His	Phe
BT:				Ile	Val	Gly	Gly	Tyr	Thr	Cys	Gly	Ala	Asn	Thr	Val	Pro	Tyr	Gln	Val	Ser	Leu	Asn			-	-	-	Ser	Gly	Tyr	His	Phe
SGT:				Val	Val	Gly	Gly	Thr	Arg	Ala	Ala	Gln	Gly	Glu	Phe	Pro	Phe	Met	Val	Arg	Leu				-	-	-	-	-	Ser	Met	Gly

	42	42A	43	44	44A	45	46	47	48	49	50	50A	50B	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	65A
α-LP:	Cys	Ser	Val	Gly	Phe	Ser	Val	Thr	Arg	Gly	Ala	Thr	Lys	Gly	Phe	Val	Thr	Ala	Gly	His	Cys	Gly	Thr	Val	Asn	Ala	Thr	Ala	Arg
PE:	Cys	-	Gly	Gly	-	Thr	Leu	Ile	Arg	Gln	Asn	-	-	Trp	Val	Met	Thr	Ala	Ala	His	Cys	Val	Asp	Arg	Glu	Leu	Thr	Phe	Arg
BCA:	Cys	-	Gly	Gly	-	Ser	Leu	Ile	Asn	Glu	Asn	-	-	Trp	Val	Val	Thr	Ala	Ala	His	Cys	Gly	Val	Thr	Thr	Ser	Asp	Val	-	
BCB:	Cys	-	Gly	Gly	-	Ser	Leu	Ile	Ser	Glu	Asp	-	-	Trp	Val	Val	Thr	Ala	Ala	His	Cys	Gly	Val	Thr	Thr	Ser	Asp	Val	-	
BT:	Cys	-	Gly	Gly	-	Ser	Leu	Ile	Asn	Ser	Gln	-	-	Trp	Val	Val	Ser	Ala	Ala	His	Cys	Tyr	Lys	Ser	Gly	Ile	Gln	Val	Arg	
SGT:	Cys	-	Gly	Gly	-	Ala	Leu	Tyr	Ala	Gln	Asp	-	-	Ile	Val	Leu	Thr	Ala	Ala	His	Cys	Val	Ser	Gly	Ser	Gly	Asn	Asn	Thr	

of their X-ray diffraction studies of DIP-inhibited bovine trypsin, with the structural information known about SGT (35). The numbering system of α-chymotrypsin will be used throughout this comparison unless designated otherwise. Reference should be made to Table XXXI for alignments of sequences of SGT and other serine proteases.

A gross examination of the bovine trypsin molecule has been described by the previous workers and indicates a molecule folded in two halves with the two termini of the molecule forming trans-molecular straps which terminate in opposite halves of the molecule. In addition, the halves of the molecule are restrained by two disulphide bridges not observed in chymotrypsin or SGT.

In looking at more detailed aspects of the structure of bovine trypsin, it should be recalled that the unique specificity for positively charged side chains was discussed at length in the general introduction when the importance of aspartic acid-189 was delineated. Several unique aspects to the trypsin specificity pocket were also indicated. For example, it was observed that the carbonyl oxygen of serine-214 was suitably oriented to form a hydrogen bond with the last N-H group on the substrate. This was proposed earlier by Steitz et al (1969) in their studies on the substrates and inhibitors bound to the active site of α-chymotrypsin (165). It is of considerable interest that this residue is conserved

in SGT. Similarly, three residues are indicated by Stroud and co-workers to be intimately associated with the construction of the rear of the specific binding pocket in trypsin and chymotrypsin (35). These are valine-213, tyrosine-228 and serine-190. Identical residues exist at all these positions in SGT, with the exception of serine-190 which is a threonine in SGT. The rear of the pocket is further constructed by a loop between position 183 and 190 which is considerably lengthened in trypsin by insertion of a lysine at 188A and a tyrosine at 184A. The latter tyrosine has been found in all trypsins subjected to sequence analysis. This is no less true of SGT. The latter enzyme has a valine instead of a lysine at position 188A and has an insertion of a proline residue at postion 184B, further increasing the length of this loop. The tyrosine residue is believed to be instrumental in insulating the specific binding pocket from the external milieu (35).

Hydrogen bonding in an antiparallel β-pleated sheet arrangement between peptide inhibitors and certain functional groups on the enzyme, has been implicated in the binding of peptides to chymotrypsin (87). The diffraction studies with bovine trypsin indicated that, through a deletion at position 218, trypsin had the possibility of forming four such hydrogen bonds in contrast to only three in chymotrypsin. This must be

an important property for trypsin molecules, as an identical deletion followed by a glycine at position 219, is observed in the bacterial enzyme, remarkably conserved from mutation by selective processes.

Leaving the specificity site, the internal residues of bovine trypsin, judged to be inaccessible to solvent molecules, have been observed on comparison with chymotrypsin and elastase to be remarkably conserved since the divergence of these enzymes from their common ancestor. That this is also true of the internal residues of chymotrypsin and SGT was indicated by the facility with which the ribbon diagram of SGT could be constructed using chymotrypsin as a reference molecule and the high degree of homology observed with these buried residues. This was also found to be true by Hartley, on construction of a theoretical trypsin and elastase model using the α-chymotrypsin electron density map (2). Indeed, Stroud et al (1971) have stated that about 60% of the internal residues in bovine trypsin, bovine chymotrypsin and porcine elastase are identical in all three enzymes (35). As a result of this extensive similarity, the later workers have turned their attention to sites of greatest difference within the hydrophobic interior of other enzymes, observed on comparison with the tertiary structure of bovine trypsin. With respect to SGT, it was concluded that the published

sequences which correspond to the internal regions of the bovine sequence, give strong evidence to suggest that these enzymes have a very similar conformation and beautifully illustrate internal compensation. The latter term refers to changes in the interior which are accompanied by other changes which in turn, counteract the first, permitting the main chain folding to remain unaltered. Of the 59 residues from the disulphide bridge peptides of SGT that were utilized in the previous comparison, only two amino acid substitutions could not readily be accommodated into the bovine tertiary structure. These were isoleucine-181 and methionine-199 which substituted for a phenylalanine and valine, respectively. However, the latter two hydrophobic residues from the bovine enzyme are in van der Waals contact with each other. In addition, the replacements in SGT are hydrophobic for hydrophobic substitutions with the less bulky isoleucine-181 being matched by a more bulky methionine-199. Stroud and co-workers suggested that this internal compensation supports the hypothesis that this enzyme has a highly homologous tertiary structure with that of bovine trypsin. This is identical to the conclusion made previously in this thesis on comparison of the complete SGT sequence with chymotrypsin. Since the previous workers stated that chymotrypsin and trypsin had very similar internal tertiary structures, it

follows that the conclusions regarding our more extensive comparative study with the complete SGT sequence and the α-chymotrypsin tertiary structure, can be enlarged to indicate that SGT is also highly homologous in tertiary structure with bovine trypsin, in agreement with the proposal of Stroud and co-workers (35).

As mentioned earlier, surface side chains lack the high degree of conservation and homology observed with the internal residues. As a result, regions of homology on the surface of the molecule might suggest that the residues involved have specific functions which are selected for during evolution. Stroud et al (1971) have observed that histidine-40, found in all mammalian serine proteases, participates in two hydrogen bonds in bovine trypsin—one to the hydroxyl oxygen of serine-32, and the other to the carbonyl of glycine-193 (35). Apparently, the latter two groups form a different hydrogen-bonded structure with the side chain of aspartic acid-194 in the chymotrypsinogen molecule (166). Since these residues are preserved in all the mammalian protease sequences, it was suggested that there might be a selective advantage in their conservation, which is perhaps associated with zymogen activation. In this respect, it should be pointed out that neither α-lytic protease nor SGT has a histidine residue at this position in agreement with the fact that no inactive precursor

molecules have been found for these bacterial enzymes.

It is interesting to note that in conjunction with zymogen activation, the activation peptide in trypsinogen, namely Val-$(Asp)_4$-Lys-, has been suggested to be ion paired with an acidic and four basic residues on the main chain of the molecule (35). Specifically, lysine-15 was suggested to pair with glutamic acid-186 drawing the other four aspartic acid residues close enough to pair with either lysine-145 or 188A and lysines at positions 15, 222, and 224. Again, it is significant that SGT has no basic residues at these positions, with the exception of the universal basic residue at position 145 which is an arginine in SGT. This once more indicates that this enzyme probably does not have a zymogen or that if a zymogen does indeed exist, that it is not structurally similar to trypsinogen in this area of the molecule.

In reference to the localization of the stabilizing calcium ions spoken of earlier, Stroud and colleagues have implicated aspartic acid-71 and glutamic acid-78 or aspartic acid-71 and 153 in ion pair cross-linking (35). It is their contention that the formation of a calcium ion charge interaction occurs, stabilizing the molecule by cross-linking two acidic residues in one of several surface loops on the molecule. Examination of the SGT sequence indicates that only aspartic acid-77

is present in the required position and acidic residues do not exist in the other positions. Nor are there adjacent acidic residues which might act as alternates in SGT. As a result, the hypothesized role for calcium ion stabilization in bovine trypsin does not appear valid for SGT.

Although the surface residues implicated in forming ion pairs with the activation peptide are absent in SGT, along with the previously mentioned residues implicated in the calcium ion binding of bovine trypsin, the overwhelming homology of other internal and even certain external residues, clearly underscores the extensive similarities between SGT and bovine trypsin.

In summary, this work has dealt with the physico-chemical and structural characterization of the bacterial enzyme Streptomyces griseus trypsin. Results obtained from this research have allowed for extensive comparisons between this enzyme and other serine proteases, with specific reference to bovine trypsin. The data compiled are virtually unanimous in their indication of the high degree of similarity between SGT and bovine trypsin, strongly suggesting that both enzymes have evolved from a common ancestoral protein. It is hoped that the elucidation of some of this enzyme's properties, presented in this thesis, will facilitate further research into the general mechanism of action of this class of proteases.

BIBLIOGRAPHY

1. Hartley, B. S. Ann. Rev. Biochem. 29:45 (1960).

2. Hartley, B. S. Phil. Trans. Roy. Soc., London, B257:77 (1970).

3. Hartley, B. S. and Kilby, B. A. Biochem. J. 56:288 (1954).

4. Linderstrøm-Lang, K. Z. Physiol. Chem. 182:151 (1929).

5. Maroux, S., Baratti, J. and Desnuelle, P. J. Biol. Chem. 246:5031 (1971).

6. Fruton, J. S. In "The Enzymes", (P. D. Boyer, ed.), 3rd edition, Vol. III, p. 119. Academic Press, New York. (1971).

7. Ryle, A. P. In "Methods in Enzymology", (G. E. Perlmann and L. Lorand, eds.), Vol. XIX, p. 316. Academic Press, New York. (1970).

8. Mycek, M. J. In "Methods in Enzymology", (G. E. Perlmann and L. Lorand, eds.), Vol. XIX, p. 285. Academic Press, New York. (1970).

9. Greenbaum, L. M. In "The Enzymes", (P. D. Boyer, ed.), 3rd edition, Vol. III, p. 475. Academic Press, New York. (1971).

10. Back, N. Fed. Am. Soc. Exp. Biol. 25:77 (1966).

11. Drenth, J., Jansonius, J. N., Koekoek, R. and Wolthers, B. G. In "The Enzymes", (P. D. Boyer, ed.), 3rd edition, Vol. III, p. 485. Academic Press, New York. (1971).

12. Markland, F. S. and Smith, E. L. In "The Enzymes", (P. D. Boyer, ed.), 3rd edition, Vol. III, p. 561. Academic Press, New York. (1971).

13. Matsubara, H. and Feder, J. In "The Enzymes", (P. D. Boyer, ed.), 3rd edition, Vol. III, p. 721. Academic Press, New York. (1971).

14. Kraut, J. In "The Enzymes", (P. D. Boyer, ed.), 3rd edition, Vol. III, p. 547. Academic Press, New York. (1971).

15. Titani, K., Hermodson, M. A., Ericsson, L. H., Walsh, K. A. and Neurath, H. Nature, 238:35 (1972).

16. Wahlby, S. and Engström, L. Biochim. Biophys. Acta, 151:402 (1968).

17. Nomoto, M. and Narahashi, Y. J. Biochem. 46:653 (1959).

18. Narahashi, Y., Shibuya, K. and Yanagita, M. J. Biochem. 64:427 (1968).

19. Narahashi, Y. and Yanagita, M. J. Biochem. 62:633 (1967).

20. Jurasek, L., Johnson, P., Olafson, R. W. and Smillie, L. B. Can. J. Biochem. 49:1195 (1971).

21. Johnson, P. and Smillie, L. B. Can. J. Biochem. 49:1083 (1971).

22. Dayhoff, M. O. (ed.) "Atlas of Protein Sequence and Structure", Vol. V, p. 57. National Biomedical Research Foundation, Maryland. (1972).

23. Bender, M. L. and Kezdy, F. J. Ann. Rev. Biochem. 34:49 (1965).

24. Blow, D. M., Birktoft, J. J. and Hartley, B. S. Nature, 221:337 (1969).

25. Zerner, B. and Bender, M. L. J. Am. Chem. Soc. 86:3669 (1964).

26. Bender, M. L. J. Am. Chem. Soc. 84:2582 (1961).

27. Bender, M. L. and Kezdy, F. J. J. Am. Chem. Soc. 86:3704 (1964).

28. Stauffer, C. E. and Zeffren, E. J. Biol. Chem. 245:3282 (1970).

29. Johnson, P. and Smillie, L. B. Can. J. Biochem. 50:591 (1972).

30. Bender, M. L., Killheffer, J. V. and Kezdy, F. J. J. Am. Chem. Soc. 86:5330 (1964).

31. Mares-Guia, M. and Shaw, E. Fed. Am. Soc. Exp. Biol. Proc. 22:528 (1963).

32. Polgar, L. and Bender, M. L. Proc. Nat. Acad. Sci. U.S.A. 64:1335 (1969).

33. Wang, J. H. Science, 161:328 (1968).

34. Shotton, D. M. and Watson, H. C. Nature, 225:811 (1970).

35. Stroud, R. M., Kay, L. M. and Dickerson, R. E. Cold Spring Harbor Symp. Quant. Biol. 36:125 (1971).

36. Walsh, K. A. In "Methods in Enzymology", (G. E. Perlmann and L. Lorand, eds.), Vol. XIX, p. 41. Academic Press, New York. (1970).

37. Yoshida, N., Sasaki, A. and Inoue, H. Fed. Eur. Bio. Soc. Letters 15:129 (1971).

38. Morihara, K. and Tsuzuki, H. Arch. Biochem. Biophys. (Comm.) 126:971 (1968).

39. Baines, N. J., Baird, J. B. and Elmore, D. T. Biochem. J. 90:470 (1964).

40. Trowbridge, C. G., Krehbiel, A. and Laskowski, M. Biochem. 2:843 (1963).

41. Chevallier, J. and Yon, J. Biochem. Biophys. Acta, 112:116 (1966).

42. Curragh, E. F. and Elmore, D. T. Biochem. J. 93:163 (1964).

43. Webster, M. E. and Prado, E. In "Methods in Enzymology", (G. E. Perlmann and L. Lorand, eds.), Vol. XIX, p. 699. Academic Press, N. Y. (1970).

44. Murachi, T. In "Methods in Enzymology", (G. E. Perlmann and L. Lorand, eds.), Vol. XIX, p. 283. Academic Press, N. Y. (1970).

45. Gorini, L. Biochem. Biophys. Acta, 7:318 (1951).

46. Sipos, T. and Merkel, J. R. Biochem. 9:2766 (1970).

47. Lazdunski, M. and Delaage, M. Biochim. Biophys. Acta, 105:541 (1965).

48. Cunningham, L. W., Tietze, F., Green, N. M. and Neurath, H. Disc. Faraday Soc. 13:58 (1953).

49. Green, N. M. and Neurath, H. In "The Proteins", (H. Neurath and K. Bailey, eds.) Vol. II B, p. 1057. Academic Press, New York (1954).

50. Travis, J. and Liener, I. E. J. Biol. Chem. 240:1962 (1965).

51. Travis, J. and Roberts, R. C. Biochem. 8:2884 (1969).

52. McDonald, M. R. and Kunitz, M. J. Gen. Physiol. 25:53 (1941).

53. Keil, B. In "The Enzymes", (P. D. Boyer, ed.), 3rd edition, Vol. III, p. 249. Academic Press, New York. (1971).

54. Abita, J. P., Delaage, M., Lazdunski, M. and Savrda, J. Eur. J. Biochem. 8:314 (1969).

55. Kraut, J. In "The Enzymes", (P. D. Boyer, ed.), 3rd edition, Vol. III, p. 165. Academic Press, New York. (1971).

56. Oppenheimer, H. L., Labouesse, B. and Hess, G. P. J. Biol. Chem. 241:2720 (1966).

57. Sigler, P. B., Blow, D. M., Matthews, B. W. and Henderson, R. J. Mol. Biol. 35:143 (1968).

58. Bergmann, M. and Fruton, J. S. J. Biol. Chem. 118:405 (1937).

59. Bergmann, M., Fruton, J. S. and Pollock, H. J. Biol. Chem. 122:643 (1939).

60. Eyl, A. and Inagami, T. Biochem. Biophys. Res. Comm. 38:149 (1970).

61. Jansen, E. F., Nutting, M. and Balls, A. J. Biol. Chem. 179:201 (1949).

62. Hartley, B. S. Nature, 201:1284 (1964).

63. Baker, B. R. "Design of Active-Site-Directed Irreversible Enzyme Inhibitors", p. 13. John Wiley and Sons, Inc., New York. (1967).

64. Fahrney, D. E. and Gold, A. M. J. Am. Chem. Soc. 85:997 (1963).

65. Bender, M. L. and Kaiser, E. T. J. Am. Chem. Soc. 84:2556 (1961).

66. Hartley, B. S. and Kilby, B. A. Biochem. J. 56:288 (1954).

67. Desnuelle, P. In "The Enzymes", (P. D. Boyer, ed.), 2nd edition, Vol. IV, p. 93. Academic Press, New York. (1960).

68. Himoe, A., Parks, P. C., and Hess, G. P. J. Biol. Chem. 242:919 (1967).

69. McConn, J., Fasman, G. D. and Hess, G. P. J. Mol. Biol. 39:551 (1969).

70. Valenzuela, P. and Bender, M. L. Proc. Nat. Acad. Sci. U.S.A. 63:1214 (1969).

71. Valenzuela, P. and Bender, M. L. Biochem. 9:2440 (1970).

72. Agarwal, S. P., Martin, C. J., Blair, T. T. and Marini, M. Biochem. Biophys. Res. Comm. 43:510 (1971).

73. Kaplan, H. and Dugas, H. Biochem. Biophys. Res. Comm. 34:681 (1969).

74. Kaplan, H., Stevenson, K. J., and Hartley, B. S. Fed. Am. Soc. Exp. Biol. Proc. 28:533 (1969).

75. Paterson, G. M. and Whitaker, D. R. Can. J. Biochem. 47:317 (1969).

76. Kaplan, H. and Whitaker, D. R. Can. J. Biochem. 47:305 (1969).

77. Labouesse, J. and Gervais, M. Eur. J. Biochem. 2:215 (1967).

78. Scrimger, S. J. and Hofmann, T. J. Biol. Chem. 242:2528 (1967).

79. Robinson, N. C., Neurath, H. and Walsh, K. A. Biochem. 12:420 (1973).

80. Hess, G. P. In "The Enzymes", (P. D. Boyer, ed.), 3rd edition, Vol. III, p. 213. Academic Press, New York. (1971).

81. Beeley, J. G. and Neurath, H. Biochem. 7:1239 (1968).

82. Blow, D. M., Birktoft, J. J. and Hartley, B. S. Nature, 221:337 (1969).

83. Smillie, L. B., Furka, A., Nagabhushan, N., Stevenson, K. J. and Parkes, C. O. Nature, 218:343 (1968).

84. Perutz, M. F., Kendrew, J. C. and Watson, H. C. J. Mol. Biol. 13:669 (1965).

85. Gutte, B. and Merrifield, R. B. J. Am. Chem. Soc. (Comm.) 91:501 (1969).

86. Olafson, R. W., Jurasek, L., Johnson, P. and Smillie, L. B. Can. Fed. Biol. Soc. Proc. 14:12 (1971).

87. Segal, D. M., Powers, J. C., Cohen, G. H., Davies, D. R. and Wilcox, P. E. Biochem. 10:3728 (1971).

88. Margoliash, E. The Harvey Lectures, Series 66, p. 177. Academic Press, New York. (1972).

89. Hartley, B. S., Brown, J. R., Kauffman, D. L. and Smillie, L. B. Nature, 207:1157 (1965).

90. Hartley, B. S. Biochem. J. (Proc.) 110:2 (1968).

91. Neurath, H., Bradshaw, R. A. and Arnon, R. In "Structure-Function Relationships of Proteolytic Enzymes", (P. Desnuelli, H. Neurath and M. Ottesen, eds.), p. 113. Academic Press, New York. (1970).

92. Nirenberg, M., Caskey, T., Marshall, R., Brimacombe, R., Kellogg, D., Doctor, B., Hatfield, D., Levin, J., Rottmann, F., Pestka, S., Wilcox, M. and Anderson, F. Cold Spring Harbor Symp. Quant. Biol. 31:11 (1966).

93. Crick, F. H. C. Symp. Soc. Exp. Biol. 12:138

94. Nishimura, S., Jones, D. S., Wells, R. D., Jacob, T. M. and Khorana, H. G. Fed. Am. Soc. Exp. Biol. Proc. 24:409 (1965).

95. Yanofsky, C., Carleton, B. C., Guest, J. R., Helinski, D. R. and Henning, U. Proc. Nat. Acad. Sci. U.S.A. 51:266 (1964).

96. Bailey, J. L. "Techniques in Protein Chemistry", Elsevier Publishing Co., Amsterdam. (1962).

97. Florkin, M. In "Taxonomic Biochemistry and Serology", (C. A. Leone, ed.), p. 51. Ronald Press Co., New York. (1964).

98. Sanders, M. M., Walsh, K. A. and Arnon, R. Biochem. 9:2356 (1970).

99. Dixon, G. H. Essays in Biochemistry 2:148 (1966).

100. Fox, S. W. (ed.) "The Origin of Prebiological Systems", Academic Press, New York. (1960).

101. Pirie, N. W., Int. Symp. Origin Life Earth 76:83 (1959).

102. Florkin, M. (ed.) "Aspects of the Origin of Life", Oxford: Permagon Press. (1960).

103. Hunt, J. A. and Ingram, V. M. Nature, 184:640 (1959).

104. Ingram, V. M. Nature, 189:704 (1961).

105. Bonner, D. M., DeMoss, J. A. and Mills, S. E. In "Evolving Genes and Proteins", (V. Bryson and H. J. Vogel, eds.), p. 305. Academic Press, New York. (1965).

106. Epstein, C. J., Nature, 203:1350 (1964).

107. Kendrew, J. C. Brookhaven Symp. Biol. 15:216 (1962).

108. Gray, W. R. Ph.D. Thesis, St. John's College, The University of Cambridge, Cambridge, England. (1964).

109. Jurasek, L., Fackre, D. and Smillie, L. B. Biochem. Biophys. Res. Comm. 37:99 (1969).

110. Landon, M. Ph.D. Thesis, University of Birmingham, England. (1964).

111. Welinder, K. and Smillie, L. B. Can. J. Biochem. 50:63 (1972).

112. Dreyer, W. J. and Bynum, E. In "Methods in Enzymology", (C. H. W. Hirs, ed.), Vol. XI, p. 32. Academic Press, New York. (1967).

113. Ryle, A. P., Sanger, F., Smith, L. F. and Katai, R. Biochem. J. 60:541 (1955).

114. Heilmann, J., Barollier, J., and Watzke, E., Hoppe-Seyl, Z. 309:219 (1957).

115. Dent, C. E. Biochem. J. 41:240 (1947).

116. Acher, R. and Crocker, C. Biochim. Biophys. Acta, 9:704 (1952).

117. Reddi, K. K. and Kodicek, E. Biochem. J. 53:286 (1953).

118. Mazur, R. H., Ellis, B. W. and Cammarata, P. S. J. Biol. Chem. 237:1619 (1962).

119. Hodges, R. A. and Smillie, L. B. Can. J. Biochem. 50:63 (1972).

120. Hartley, B. S. Biochem. J. 119:805 (1970).

121. Sodek, J. Unpublished.

122. Gray, W. R. In "Methods in Enzymology", (C. H. W. Hirs, ed.), Vol. XI, p. 469. Academic Press, New York. (1967).

123. Chervenka, C. H. "A Manual of Methods for the Analytical Ultracentrifuge", Spinco Division of Beckman Instruments, Publishers. Palo Alto, California (1970).

124. Richards, E. S. and Schachman, H. K. J. Phys. Chem. 63:1578 (1959).

125. Kay, C. M. and Oikawa, K. Biochem. 5:213 (1966).

126. Oikawa, K., Kay, C. M. and McCubbin, W. D. Biochim. Biophys. Acta, 168:164 (1968).

127. Moffit, W. and Yang, J. T. Proc. Nat. Acad. Sci. U.S.A. 42:590 (1956).

128. Todd, A. In "A Laboratory Manual of Analytical Methods of Protein Chemistry Including Polypeptides", (P. Alexander and R. J. Block, eds.), Vol. II, p. 245. Permagon Press, Oxford. (1960).

129. Jirgensons, B. In "Optical Rotatory Dispersion of Proteins and Other Macromolecules", Springer-Verlag, New York, p. 48. (1969).

130. Greenfield, N. and Fasman, G. D. Biochem. 8:4108 (1969).

131. Schwert, G. W. and Takenaka, Y. Biochim. Biophys. Acta, 16:570 (1955).

132. Walsh, K. A. and Wilcox, P. E. In "Methods in Enzymology", (G. E. Perlmann and L. Lorand, eds.), Vol. XIX, p. 31. Academic Press, New York. (1970).

133. Glasoe, P. K. and Long, F. A. J. Phys. Chem. 64:188 (1960).

134. Wahlby, S. Biochim. Biophys. Acta, 151:394 (1968).

135. Moore, S. and Stein, W. H. J. Biol. Chem. 211:907 (1954).

136. Gross, E. and Witkop, B. J. Biol. Chem. 237:1856 (1962).

137. Raftery, M. A. and Cole, R. D. J. Biol. Chem. 241:3457 (1966).

138. Butler, P. J. G., Harris, J. I., Hartley, B. S. and Leberman, R. Biochem. J. 112:679 (1969).

139. Stark, G. R., Stein, W. H. and Moore, S. J. Biol. Chem. 235:3177 (1960).

140. Rimon, S. and Perlmann, G. J. Biol. Chem. 243:3566 (1968).

141. Hirs, C. H. W. In "Methods in Enzymology", (C. H. W. Hirs, ed.), Vol. XI, p. 197. Academic Press, New York. (1967).

142. Shaw, E. In "Methods in Enzymology", (C. H. W. Hirs, ed.), Vol. XI, p. 677. Academic Press, New York. (1967).

143. Brown, J. R. and Hartley, B. S. Biochem. J. 89:59 P (1963).

144. Stevenson, K. J. and Smillie, L. B. Can. J. Biochem. 46:1357 (1968).

145. Edman, P. and Begg, G. Eur. J. Biochem. 1:80 (1967).

146. Yamada, S. and Itano, H. Biochem. Biophys. Acta, 130:538 (1966).

147. Butler, J. N. In "Solubility and pH Calculations", Addison-Wesley Publishing Company, Reading, Massachusetts. (1964).

148. Weber, K. and Osborn, M. J. Biol. Chem. 244:4406 (1969).

149. Lineweaver, H. and Burk, D. J. Amer. Chem. Soc. 56:658 (1934).

150. Gutfreund, H. Disc. Farad. Soc. 20:167 (1955).

151. Gutfreund, H. Disc. Farad. Soc. 51:441 (1955).

152. Mahler, H. R. and Cordes, E. H. In "Biological Chemistry", p. 10. Harper and Row, Publishers, New York. (1966).

153. McCubbin, W. D., Kay, C. M. and Oikawa, K. Biochim. Biophys. Acta, 120:597 (1966).

154. Jirgensons, B. In "Optical Rotatory Dispersion of Proteins and Other Macromolecules", Springer-Verlag, New York, p. 15. (1969).

155. Narahashi, Y. and Fukunaga, J. J. Biochem. 66:743 (1969).

156. Trop, M. and Birk, Y. Biochem. J. 116:19 (1970).

157. Schoellman, G. and Shaw, E. Biochem. Biophys. Res. Comm. 7:36 (1962).

158. Mares-Guia, M. and Shaw, E. J. Biol. Chem. 240:1579 (1965).

159. Kenner, G. W. J. Chem. Soc. D 636 (1971).

160. Bender, M. L. and Hamilton, G. A. J. Am. Chem. Soc. 84:2570 (1962).

161. Bender, M. L., Pollock, E. J. and Neveu, M. C. J. Am. Chem. Soc. 84:595 (1962).

162. Robinson, N. C., Ross, W. T., Neurath, H. and Walsh, K. A. Biochem. 10:2743 (1971).

163. Dixon, J. W. and Hofmann, T. Can. J. Biochem. 48:671 (1970).

164. Blow, D. M. and Steitz, T. A. Ann. Rev. Biochem. 39:63 (1970).

165. Steitz, T. A., Henderson, R. and Blow, D. M. J. Mol. Biol. 46:337 (1969).

166. Ferer, S. T., Kraut, J., Robertus, J. D., Wright, H. T. and Xuong, N. H. Biochem. 9:1997 (1970).

167. Jurasek, L. and Smillie, L. B. Can. J. Biochem. (In press).

168. Siegel, S., Brady, A. H. and Awad, W. M. J. Biol. Chem. 247:4155 (1972).

169. Schroeder, W. A., Shelton, J. R. and Robberson, B. Biochim. Biophys. Acta, 147:590 (1967).

170. Offord, R. E. Nature, 211:591 (1966).

171. Haley, E. E. and Corcoran, B. J. Biochem. 6:2668 (1967).

172. Smyth, D. G., Stein, W. H. and Moore, S. J. Biol. Chem. 237:1845 (1962).

173. Swallow, D. L. and Abraham, E. P. Biochem. J. 70:364 (1959).

174. Smyth, D. G., Stein, W. H. and Moore, S. J. Biol. Chem. 238:227 (1963).